

Tartu Ülikool
Loodus- ja tehnoloogiateaduskond
Ökoloogia ja Maateaduste instituut
Geograafia osakond

Magistritöö geoinformaatikas

**Maastikumustri indikaatorite tundlikkus satelliidipiltide
rasterandmestiku töötlustehnikate suhtes**

Kirke Narusk

Juhendaja: prof. Tõnu Oja

Kaitsmisele lubatud:

Juhendaja: /allkiri, kuupäev/

Osakonna juhataja: /allkiri, kuupäev/

Tartu 2014

Sisukord

Sissejuhatus.....	3
1. Teoreetiline taust.....	4
1.1. Maastikumeetrika.....	4
1.2. Probleemid maastiku indikaatorite kasutamisel	5
1.2.1. Indikaatorite ebasobivusest tingitud probleemid maastikumustri kirjeldamisel	6
1.2.2. Ebamäärasuse põhjused maastikumeetrika arvutamisel satelliidipiltidelt.....	7
2. Andmed ja metoodika.....	11
2.1. Satelliidipildid.....	11
2.2. Tarkvarad	13
2.3. Etalonideta klassifitseerimine.....	13
2.4. Etalonidega klassifitseerimine	16
2.5. Objektorienteeritud klassifitseerimine	17
2.6. Maastikuindeksite arvutamine tarkvaraga Fragstats	18
2.7. Klassifikatsioonide ja indekse võrdlemine	19
3. Tulemused.....	20
3.1. Etalonideta klassifitseerimine.....	20
3.2. Etalonidega klassifitseerimine	24
3.3. Objektorienteeritud klassifitseerimine	27
3.4. Klassifitseerimistulemuste põhjal arvutatud maastikuindeksite varieerumine.....	28
4. Arutelu	41
4.1. Tarkvaratöötlus	41
4.2. Maastikuindeksid	43
Kokkuvõte.....	46
Summary	47
Kasutatud kirjandus.....	49
Lisad	53

Sissejuhatus

Maastikumeetrikat ehk maastikuindekseid kasutatakse maastikuökoloogias mõõtmaks maastiku struktuuri. Struktuuril on tugev seos bioloogilise mitmekesisuse ning ökoloogiliste protsesside ja toimemehhanismidega, mistõttu maastikumeetrika on oluline vahend maastikuplaneerimise jaoks (Turner 1989; McGarigal & Marks 1995; Leitão *et al.* 2006).

Satelliidipilte ja nende töötlemiseks sobivaid tarkvarapakette kasutatakse laialdaselt maastikuindeksite arvutamiseks, kuid sageli ei hinnata satelliidipiltide omadustest või tarkvaratöötlustest tulenevat ebamäärasust indeksite väärtustele (Fassnacht *et al.* 2006; Lechner *et al.* 2012). Satelliidipiltide omadused ja töötlus mõjutavad aga maastikustruktuuri analüüsi tulemusi ja on tugevas seoses enamiku indeksite väärtustega (Turner 1989; McGarigal 2002; Baldwin 2004). Kuigi satelliidipiltide kasutamisel maastikumeetrika arvutamiseks on peamiseks puuduseks klassifitseerimise suhteline keerukus, siis eeliseks on võimalus jälgida ajalist muutust, mida olemasolevad vektorkaardid nii hästi ei võimalda (Uemaa *et al.* 2008). Satelliidipiltide mõõtkavast, sh ruumilisest lahutusest tulenevat mõju maastikumeetrikale on üsna palju uuritud (Saura 2004; Wu 2004; Símová, Gdulová 2012), uurimusi on ka klassifitseerimise skeemi (Baldwin *et al.* 2004; Buyantuyev, Wu 2007; Peng *et al.* 2007) ja klassifitseerimisvigade (Shao *et al.* 2001; Langford *et al.* 2006; Altamirano *et al.* 2012) mõju kohta, kuid mil määral mõjutavad eri tarkvarade töötlustehnikad ja algoritmid indeksite väärtust, on saanud vähem tähelepanu. Satelliidipiltide klassifitseerimis- ja töötlusmeetodite mõju on uurinud näiteks Mas *et al.* (2010), Lechner *et al.* (2013), kuid üldiselt on sellekohaseid uuringuid vähe.

Maastikuindekseid on kasutusel väga palju, rohkem, kui neid enamasti maastiku kirjeldamiseks tarvis läheb, mistõttu erinevad uurijad on püüdnud leida indeksite tuumikut, mis kirjeldaksid kõige paremini maastiku eri aspekte (Riitters *et al.* 1995; Lausch, Herzog 2002; Leitão *et al.* 2006; Uemaa *et al.* 2011). Oluline on aga ka see, kas need indeksid on tundlikud tehnilise töötluste suhtes.

Kuna maastikuindeksite väärtusi mõjutavad nii satelliidipiltide töötlusmeetodid kui maastikumuster (Lechner *et al.* 2013), siis mustri paremaks interpreteerimiseks oleks vajalik mõista töötluste mõju indeksite arvutustele. Oluline on seejuures ka tegeliku maastikumustri ja töötlusmeetodite suhe: kas üks või teine töötlemise viis mõjutab indeksi väärtust kõigil maastikel ühtmoodi või sõltub mõju maastikutüübist. Käesoleva **magistritöö eesmärk** on uurida, kui palju mõjutab satelliidipiltide tarkvaratöötlust maastikumeetrika arvutamist ja millised maastikku iseloomustavad indeksid on töötlustest kõige vähem sõltuvad. Selleks võrreldakse nelja kaugseires kasutatava kommerts-pilditöötlustarkvara pakutavaid töötlustehnikaid suure ruumilise lahutusega satelliidipiltide peal. Tulemuseks saadakse eri töötlusviisidest tulenevad indeksite varieeruvused, millega saab hinnata töötluste mõju indeksite väärtustele. See informatsioon aitab maastikuökolooge ja -planeerijaid sobivate indeksite valikul.

1. Teoreetiline taust

1.1. Maastikumeetrika

Maastikuindeksid on algoritmid maastiku geomeetria ja ruumiliste omaduste kvantitatiivseks kirjeldamiseks teatud skaalas (mõõtkavas ja ulatuses) (McGarigal 2002). Neid kasutatakse maastikuökoloogias uurimaks maastikumustrit ja selle mõju ökosüsteemidele, kuidas muster mõjutab protsesse ja protsessid mustrit (Turner 1989). Maastikumuster tekib füüsikaliste, bioloogiliste ja sotsiaalsete tegurite koosmõjul. Konfiguratsioon (eraldiste kuju, suurus, paiknemine, tihedus, fraktaalne dimensioon, koonduvus) ja kompositsioon (mitmekesisus, proportsioonid, klasside arv) on maastikumustri kaks peamist komponenti (McGarigal & Marks 1995). Alates 1980. aastatest on välja pakutud üle saja indeksi, analüüsivaks maastiku kompositsiooni ja konfiguratsiooni. Enamikku neist indeksitest saab arvutada programmi Fragstats abil (McGarigal & Marks 1995). Kõige enam on indekseid kasutatud maakasutuse ja –katte muutuste hindamiseks, aga ka elupaikade ja bioloogilise mitmekesisuse, regulatsioonifunktsioonide (kliima, mikrokliima, veekvaliteet, üleujutuste kontroll, tulekahjude leviku kontroll) ning vähemal määral ka maastiku esteetika ja tajumise uurimiseks (Uuemaa *et al.* 2012).

Maastikuindekseid saab rühmitada taseme ning indeksi tüübi alusel (McGarigal 2002). Eristatakse kolme indeksite arvutamise taset: eraldis, klass ning maastik. Eraldiseks loetakse diskreetset maastiku elementi, mille struktuur või kompositsioon erineb ümbritsevast alast. Eraldise taseme indeksid iseloomustavad üksikeraldise ruumilist iseloomu ja konteksti (McGarigal 2002). Klassi taseme indeksid arvutatakse kõigi sama tüüpi eraldiste kohta ja need näitavad eraldiste klassi ruumilist paigutust, osakaalu maastikus ja sobivad seetõttu fragmenteerituse kirjeldamiseks (McGarigal *et al.* 2012). Maastiku taseme indeksid kirjeldavad kogu maastiku – kõikide eraldiste ja eraldiste klasside ruumilist jaotust ning kompositsiooni. Neid saab kasutada maastiku mitmekesisuse hindamiseks (McGarigal *et al.* 2012). Kõiki indekseid ei ole võimalik arvutada kolmel tasemel, kuid enamikul eraldise taseme indeksitel on vasted klassi ja maastiku tasemel. Paljud maastiku ja klassi taseme indeksid arvutatakse eraldiste või klasside statistiliste andmete summeerimisel või keskmistamisel, mistõttu suur osa indekseid korreleeruvad omavahel (McGarigal, Marks 1995; McGarigal 2002). Indeksi tüübi alusel saab maastikumeetrikat jaotada mitmeti, sest paljusid indekseid saab määrata rohkem kui ühte kategooriasse, kuid programmis Fragstats on seda tehtud järgnevalt (McGarigal *et al.* 2012):

- Pindala ja serva indeksid (*Area & edge metrics*) – kirjeldavad eraldiste suurust, servapikkust. Olulised pindala-tundlike liikide seisukohalt. Servapikkus annab informatsiooni fragmenteerituse kohta. Sõltuvad eelkõige ruumilisest lahtusest ja määratud väikseimast eraldise suurusel / kaardistusühikust (MMU, ingl *minimum mapping unit*) ning maastiku ulatusest.
- Kuju indeksid (*Shape metrics*) – kirjeldavad eraldiste kuju ja suurust. Enamik kujuindekseid põhinevad übermõõdu-pindala suhtel. Kuju mõjutab eraldiste sees toimuvaid protsesse, nagu pisiimetajate ränne. Kuju indeksid on eriti olulised servaeefekti tõttu.
- Tuumala indeksid (*Core area metrics*) – tuumalaks loetakse ala eraldise sees, mis jääb välja servaeefekti mõjualast. Viimase määrab kasutaja vastavalt

uuritavale protsessile. Sobivad tuumalasiid servadele eelistavate liikide elupaikade kvaliteedi hindamiseks.

- Kontrastsuse indeksid (*Contrast metrics*) – kontrastus viitab külgnevate eraldiste ökoloogilistest tunnustest tulenevatele erinevustele, mis võivad mõjutada olulisi ökoloogilisi protsesse. Mida erinevamad on kõrvuti paiknevad eraldised, seda suurem kontrastsus. Kontrastsusest sõltub ka servaeefekti mõju.
- Koonduvuse indeksid (*Aggregation metrics*) – need indeksid näitavad eraldiste hajutatust / koonduvust / vahelduvust, mõõtes ühte tüüpi pikslite kõrvuti asetsemist. Hajuvus näitab ühte tüüpi eraldiste ruumilist jaotust ja nende sidusust, vahelduvus eri tüüpi eraldiste ruumilist segunemist. Sidusus iseloomustab, kui võrd organismid saavad liikuda ühest eraldisest teise, väikese sidususe korral muutuvad osad populatsioonid isoleerituks, mis võib viia populatsiooni välja suremiseni. Eraldiste koonduvus on oluline elupaikade fragmenteerituse seisukohast, iseloomustades nende isoleeritust ning liigendatust.
- Mitmekesisuse indeksid (*Diversity metrics*) – sõltuvad rohkusest (*richness*) ja ühtlusest (*evenness*). Rohkus viitab eraldiste tüüpide arvule maastikus ja ühtlus pindalalisele jaotusele.

Eristatakse veel struktuuri indekseid ja funktsionaalseid indekseid. Esimesed mõõdavad lihtsalt kompositsiooni ja konfiguratsiooni, viitamata konkreetsele ökoloogilisele protsessile, teised aga on seotud uuritava organismi või protsessiga (McGarigal 2002).

Töös kasutatavate indeksite täielik ülevaade on esitatud lisas 1.

1.2. Probleemid maastiku indikaatorite kasutamisel

Maastikumeetrikat kasutatakse rohkelt maastiku kirjeldamiseks, kuid nii mõnigi kord valesti. Näiteks ei uurita, selgitata piisavalt maastikumustri ja protsesside vahelisi seoseid või indeksite ning kaardiandmestiku asjakohasust, mis võib viia valede või mitte midagi tähendavate tulemusteni (Li, Wu 2004). Indeksiste valikul on oluline võtta arvesse uurimisala maastiku iseloomu ja uurimise eesmärki. Li ja Wu (2004) järgi esineb kolme põhilist tüüpi vigu maastikuindeksite kasutamisel: kontseptuaalsed vead (põhjendamatud seosed maastikumustri ja protsesside vahel, indeksite vähene asjakohasus protsesside kirjeldamisel, vale mõõtkava), indeksite vale kasutus (konkreetses maastiku jaoks sobimatu, puudulikud korrelatsioonianalüüsid maastiku ja indikaatorite vahel), maastikuindeksite omastest vigadest, nagu vähene tundlikkus ja üksteisega sarnanemine, tingitud kasutamise piirangud, mis raskendavad nende interpreteerimist. Sageli jääb selgusetuks, mida üks või teine indeks täpselt mõõdab, kuna mitmed tegurid võivad korraga indeksi väärtust mõjutada.

Maastikumeetrika arvutamisel hinnatakse ja esitatakse harva ruumilist ebamäärasust (*uncertainty*) (Lechner *et al.* 2012). Maastikumeetrikaga seotud ebamäärasus hõlmab täpsushinnanguid (arvutatud ja tegelike väärtuste ühtimine), tabavushinnanguid (kui lähedased on korduvmõõdistuste tulemused), tähenduslikkust (kuidas indeksite väärtuste erinevusi tõlgendada) (Brown *et al.* 2004). Need sõltuvad mitmest tegurist, nagu maastikuklasside defineerimine, klassifitseerimise viis, kaardi täpsus (klassifitseerimisvead) ja sobivus, indeksite sõltuvus mõõtkavast ning kasutatavate indeksite unikaalsus (Brown *et al.* 2004; Lechner *et al.* 2012). Maastikumustri kvantitatiivsel mõõtmisel on oluline silmas pidada, kuidas maastik on klassifitseeritud

ehk millised kategooriad on loodud ja milliseid klassifitseerimise viise kasutatud, milline on ruumiline ulatus ehk uuritava ala suurus ja mõõtkava ehk ruumiline lahutus ning andmete formaat (raster- või vektorandmed) – need mõjutavad maastikustruktuuri analüüsi tulemusi ja on tugevas seoses enamiku indeksite väärtustega (Turner 1989; McGarigal 2002; Baldwin 2004). Indikaatorite tundlikkus kaardi omaduste suhtes võib varieeruda laias ulatuses nii ruumiliselt kui indeksite endi hulgas (Baldwin *et al.* 2004).

Mõõtkava, loodud maastiku kategooriad ja ulatus peaksid olema vastavuses uuritavate protsessidega ning põhjendatud. Lechner *et al.* (2012) uurisid ajakirja Landscape Ecology maastikumeetrikat käsitlevaid artikleid ja leidsid, et neist paljudes ei mainita kasutatavate ruumiandmete päritolu ja töötlemismeetodeid, umbes pooled ei taga andmeid mõõtkavast põhjustatud ebamäärasuse ja 81% klassifitseerimisvigade kohta. Ka Uemaa *et al.* (2011) leiavad, et pole haruldane, et uurimistöodes ei mainita maakasutuse/-katte klasside arvu või kasutatud piltide ruumilist lahutust, mis maakasutuse muutuste uurimise seisukohalt on väga olulised. Lisaks võib ruumiline lahutus mõnikord osutada uuritava protsessi kirjeldamiseks ebasobivaks: ühes mõõtkavas olulised või ennustatavad protsessid ja parameetrid ei pruugi seda olla mõnes teises mõõtkavas (Turner 1989). Vahel ei piisa aga indeksite arvutamisest vaid ühes mõõtkavas, sest uuritav nähtus toimib mitmel ruumilisel tasemel (Lechner *et al.* 2012). Maastikuökoloogias kasutatakse maastiku geograafiliseks esitamiseks peamiselt diskreetseid maakatte eraldisi (Lechner *et al.* 2012). Diskreetsed klassifitseerimisskeemid põhjustavad suurema tõenäosusega ruumilist määramatust, sest sageli ei ole maastikueraldiste vahel selgeid piire ja klassifitseerida saab lugematul arvul viisidel (Lechner *et al.* 2012). Klassifitseerimisel tekkivad vead mõjutavad omakorda maastikuindeksite arvutamist: maastikuindeksite täpsus on sageli eksponentsiaalses sõltuvuses klassifitseerimistäpsusest (Langford *et al.* 2006; Shao, Wu 2008).

1.2.1. Indikaatorite ebasobivusest tingitud probleemid maastikumustri kirjeldamisel

Indeksite väärtuste ja ökoloogiliste protsesside vaheliste seoste kindlaks tegemine on maastikumeetrika üks oluline uurimisvaldkond, mõistmaks, mida indeksid tähendavad ja kuidas nad muutuvad. Seda on lihtsam teha, kui indeksil on järgmised omadused (Dramstad 2009):

- kergesti interpreteeritav
- andmed on kättesaadavad
- teaduslik alus
- asjakohane
- informatiivne
- läbipaistev, selge
- stabiilne, tundlik

Hea indeks peab andma selget informatsiooni poliitikutele ja planeerijatele, olema asjakohane maastikuplaneerimise seisukohalt. Selle eelduseks on indikaatori arvutuslik läbipaistvus, parameetrite selgus, mis indeksi väärtust mõjutavad, tundlikkus maastiku muutustele ja stabiilsus (väike varieeruvus häiringute ning muutuste suhtes). Indeksi arvutamiseks vajalikud andmed peaksid olema kergesti kättesaadavad ja asjakohased.

Ometi võivad mõned indeksid käituda ettearvamatult ja anda näiteks erinevate maastikumustrite korral samu tulemusi või vastupidi, anda samade muutuste korral erinevaid tulemusi, mistõttu nende tundlikkus on küsitav (Dramstad 2009). Lihtsam on interpreteerida indekseid, mis kirjeldavad maastiku ühte aspekti, mitte mitut, ja mille väärtus jääb kindlasse vahemikku (näiteks 0 kuni 1) (Li, Wu 2004).

Üldiselt ei soovitata kaasata uurimusse liiga palju indekseid, kuna maastiku kirjeldamiseks piisab enamasti mõnest. Paljud indeksid korreleeruvad omavahel ehk mõõdavad samasuguseid või sarnaseid maastikumustri aspekte (McGarigal, Marks 1995; McGarigal 2002). Põhjuseks on see, et indekseid on tunduvalt rohkem kui parameetreid, millel nad põhinevad. Erinevates uurimustes on püütud kindlaks teha, millised sõltumatud indeksid kirjeldavad kõige paremini maastikumustri peamisi komponente või millised on ühe või teise piirkonna maastiku kirjeldamiseks sobivaimad (McGarigal 2002; Uemaa *et al.* 2009). Kirjanduse ja ekspertide arvamuse põhjal valisid Leitão *et al.* (2006) välja kümme tuumik-maastikuindeksit, mis peaksid vastama planeerijate ja maastiku analüüsijate põhilistele vajadustele: eraldiste rohkus PR, maakatte tüübi ehk klassi osakaal maastikus CAP, eraldiste arv NP / eraldiste tihedus PD, eraldiste suurus AREA (sh keskmine AREA_MN ja pindalaga kaalutud keskmine AREA_AM pindala), eraldiste kuju indeks SHAPE, serva kontrastsus ECON, eraldiste kompaktsuse indeks GYRATE, eukleidiline lähima naabri kaugus ENN, läheduse indeks PROX, koonduvuse indeks CONTAG. Sobivate indeksite kindlaks tegemisel võib kasutada näiteks faktoranalüüsi (Riitters *et al.* 1995; Lausch, Herzog 2002; Uemaa *et al.* 2011). Uemaa *et al.* (2011) proovisid määrata sobivad indeksid Eesti maastiku kirjeldamiseks, kasutades peakomponentanalüüsi, ja leidsid, et 15 indeksi hulgast piisas vaid neljast, et kirjeldada maastiku ruumilise struktuuri kõige olulisemaid aspekte: domineerivust, kontrastsust, eraldiste kuju keerukust ja kompositsiooni.

1.2.2. Ebamäärasuse põhjused maastikumeetrika arvutamisel satelliidipiltidelt

Kaugseire andmeid ja nende töötlust GIS (geoinfosüsteem) abil kasutatakse kõige enam, kirjeldamaks maakatet ja maastikumustri ning ökoloogiliste protsesside vahelisi seoseid (Lechner *et al.* 2012). Lechner *et al.* 2012) järgi kasutasid enamik ajavahemikus 2000-2010 tehtud maakasutuse/-katte muutuste uurimustest maastiku indeksite arvutamiseks kaugseire andmeid. Kaugseire kasutamine keskkonna uurimustes on kasulik ja põhjendatud, kui vajatakse täielikku, ühtlast ruumilist katvust suuremate alade kohta, korduvmõõdistuste ja ajaliste erinevuste ning raskesti ligipääsetavate või tundlike alade seireks (Fassnacht *et al.* 2006). Kaugseire andmete põhjal arvutatud maastikumeetrikaga on seotud aga ebamäärasus, mis tuleneb klassifitseerimise viisist, ruumilisest mõõtkavast ning klassifitseerimisvigadest (Lechner *et al.* 2012). Seetõttu ei saa kaugseire andmete alusel tehtud kaarte võtta absoluutse tõena.

Mõõtkava

Ruumiline mõõtkava hõlmab piksli suurust, väikseimat kaardistusühikut (MMU – *minimum mapping unit*), ulatust, silumist ning temaatilist resolutsiooni (Lechner *et al.* 2012; Símová, Gdulová 2012). Kaardi mõõtkava muutused võivad mõjutada maastikumeetrikat kolmel juhul: resolutsiooni, ulatuse ja nii resolutsiooni kui ulatuse muutmisel. Mõõtkava muutumisest tingituna jagunevad maastikuindeksid laias laastus kahte kategooriasse: esimestel on suhe mõõtkava muutusega, teiste käitumist ei saa ette

ennustada, mis viitab nende suurele tundlikkusele teatud maastikumustrite suhtes (Wu 2004). Suhe võib olla nii astmeline, logaritmiline kui lineaarne, aga ka kõikuv / muutlik. Mõõtkava mõju maastikuindeksite käitumisele on üsna palju uuritud nii ehtsate kui simuleeritud maastike peal, kuigi autorid ei ole tulemuste osas alati ühel meelel (Símová, Gdulová 2012).

Saura ja Martinez-Millan (2001) järgi varieerub eraldiste kuju kirjeldavate indeksite väärtus märgatavalt ulatuse muutmisel, isegi kui maastikuklasside vahекord ja maastiku fragmenteeritus jäävad samaks. Lisaks sõltub indeksite varieeruvus ulatuse muutmisel loomulikult väga palju ka konkreetsest mustrist ja selle ruumilisest paiknemisest ning ulatuse muutmise viisist (ala diagonaalne ehk ühest nurgast kasvatamine, kõikidest nurkadest kasvatamine, külgnevad ruudud, liikuv aken). Keeruline on kindlaks teha, kas tundlikkus ulatuse muutuste suhtes tuleneb indeksist endast või maastiku konfiguratsioonist (Baldwin *et al.* 2004). Wu (2004) uurimusest selgub, et üldiselt on ulatuse muutmisel vähem etteaimatav mõju indeksite käitumisele, kui piksli suurusega mängimisel, sest tugevat või püsivat suhet näitas viimase puhul rohkem indekseid ning nad olid vähem muutlikud. Ruumilise lahutuse mõju indeksite väärtustele uurivates töodes tõdetakse, et oluline roll maastikumustri muutuste seisukohalt on ka pikslite liitmise tehnikal / algoritmil, mida sellistes uurimustes tihti kasutatakse (Baldwin *et al.* 2004; Saura 2004; Lechner *et al.* 2013). Võrreldes klassi tasandit maastiku tasandiga, siis on uurijad leidnud, et suhe mõõtkavaga on klassi tasandil üldiselt muutlikum (Wu 2004; Kelly *et al.* 2010). Arvestama peaks sellega, et resolutsiooni võib vähendada sateliidipiltide eel- ja järeltöötlemine, näiteks siluvate filtrite kasutamine, pikslite koondamine, pildi kaardivõrguga sidumine ja koordinaatsüsteemi muutmine (Lechner *et al.* 2012).

Ruumiline resolutsioon (piksli suurus) on oluline klassifitseerimist mõjutav tegur, määrates väikseima kaardistusühiku, mistõttu ruumilise lahutuse valimisel tuleks arvestada maakatteklassidega ja objektidega, mida mõõta tahetakse. Väike MMU põhjustab suurema eraldiste varieeruvuse, kuid huvipakkuvad eraldised võivad tagaplaanile jääda, suure MMU puhul aga väheneb klassifitseerimistäpsus ning huvipakkuvad eraldised võivad teiste klassidega kokku sulada, kombineeruda (Fassnacht *et al.* 2006). Suure resolutsiooni ja väikese MMU kombineerimisel võib tekkida aga müra: suurem servatihedus ja varieeruvus, väikesed eraldised.

Ruumilisest lahutusest sõltub temaatiline resolutsioon – kui palju ja millised klassid defineeritakse. Lineaarsed objektid, nagu kitsas oja, ei pruugi avalduda, kui ruumiline lahutus ei ole piisavalt suur (Lausch, Herzog 2002). Temaatilist resolutsiooni võib mõjutada ka ulatus. Mida suurem ulatus, seda suurem tõenäosus kaardistada haruldasi klasse (Fassnacht *et al.* 2006).

Maastikumeetrika arvutamisel võib olulist rolli mängida ja maastikust mitte tulenevaid erinevusi põhjustada ka ajaline resolutsioon ehk aasta-ajalised erinevused, sealhulgas atmosfääri tingimused ja taimestiku erinev spektraalne tundlikkus eri aastaaegadel (Brown *et al.* 2004).

Klassifitseerimine

Kõige enam kasutatakse maastikumeetrika arvutamiseks kateoorilist pinda ehk diskreetseid maakatte eraldisi, mille alusel iseloomustatakse maastiku geomeetriat ja ruumilisi omadusi teatud mõõtkavas ja ulatuses (McGarigal 2002). Teemakaartide saamiseks töödeldakse / klassifitseeritakse / üldistatakse ruumiandmeid eelnevalt, mis mõjutab maastikumeetrika arvutusi (Lechner *et al.* 2013). Oluline on aru saada, kui

suur mõju on pilditöötlusel tõelisele maastikumustrile ja kas see on ühesugune erinevate maastike puhul. Seetõttu on klassifitseerimise skeem, pilditöötluse üks olulisemaid osi maastikumeetrika seisukohalt, oluline kaugseirepiltide aina laialdasemas kasutamise valguses indeksite arvutamiseks.

Satelliidipildi klassifitseerimiseks on mitmeid viise (Jensen *et al.* 2009). Esmalt saab valida jäikade või hägusate klassifikaatorite vahel. Esimeste puhul on iga piksel määratud kindlasse klassi. Hägusaid klassifikaatoreid saab seostada aga segupikslitega, kus pikslite piires eristatakse mitu erinevate osakaaludega klassi. Mõlemal juhul peaksid klassid olema piisavalt täpsed, et kirjeldada maastikku, kuid mitte liiga detailselt määratletud, et neid oleks spektraalselt üksteisest võimalik eristada. Seejärel saab kasutaja valida näidistega ja näidisteta klassifitseerimise vahel. Näidisteta ehk õpetava valimita klassifitseerimisel jagatakse sarnaste spektraalväärtustega pikslid klastritesse, millele kasutaja annab hiljem sisulise tähenduse. Näidistega klassifitseerimisel on eesmärgiks iga piksel mõnda juba eelnevalt defineeritud klassi paigutada. Edasi tuleb määrata klassifitseerimisalgoritm. Klassifitseerimise algoritmid töötavad kas piksel-haaval või siduspikslite rühmade ehk objektide klassifitseerimise põhiselt (ingl *object-oriented classifiers*) ning põhinevad kas parameetrilistel statistikutel (eeldavad normaaljaotusega andmestikku), mitteparameetrilistel (ei eelda normaaljaotust) või mittemeetrilistel (töötavad nii numbriliste- kui ka nominaalväärtustega). Kõigest sellest tulenevalt ei pruugi erinevad klassifitseerimise viisid viia samale lõpptulemusele.

Indeksite väärtused sõltuvad maastikukategooriate arvust, mitmeks klassiks maastik jagatakse. On leitud, et paljud indeksid on tundlikud muutustele temaatilises resolutsioonis ehk klasside arvus (Baldwin *et al.* 2004; Buyantuyev, Wu 2007; Peng *et al.* 2007). Tundlikud on nii kompositsiooni kui konfiguratsiooni kirjeldavad indeksid (Buyantuyev, Wu 2007). Kaugseire piltide klassifitseerimise hõlbustamiseks on loodud standardkategooriate süsteem ehk maakatte ja maakasutuse klassifitseerimise skeem LULC (*Land Use and Land Cover classification system*), kus defineeritakse maastikukategooriad ja nende alamkategooriad (Anderson *et al.* 1976). Sama nimega klassi definitsioonid võivad siiski eri allikate järgi erineda, maastikumeetrika on tundlik nende erinevuste suhtes (Brown *et al.* 2004; Lechner *et al.* 2012). Erinevusi võib põhustada ka see, kas kasutatakse maakasutuse- või maakattepõhist määratlust.

Klassifitseerimisvead

Veahinnangud on vajalikud kaugseireandmetest koostatud kaartide asukohatäpsuse ja klassifitseerimis- ehk temaatilise täpsuse hindamiseks. Rusikareegel ütleb, et enamik maastikuindekseid saavutavad stabiilsed väärtused klassifitseerimistäpsuse juures, mis on üle 90% (Shao *et al.* 2001). Enamiku kättesaadava maakatte andmestiku klassifitseerimistäpsus jääb aga alla 85% - norm, mille Ameerika Ühendriikide Geoloogiateenistus (*United States Geological Survey*, USGS) esialgselt kehtetas. Foody (2008) peab 85% õigsust aga ebaõiglaselt väikeseks, kuna see seati algselt ulatuslike maakatteklasside kaardistamisele väikese ruumilise lahutusega (80 m) satelliidipiltidelt. Seda täpsust on püütud klassifitseerimismeetodite arendamisega parandada, kuid olulist progressi pole toimunud ja kaugseireandmetest koostatud teemakaartide täpsust peetakse seetõttu sageli küsitavaks.

Kuna maastikumeetrika väärtused on eksponentsiaalses sõltuvuses klassifitseerimistäpsusest, on oluline teada kasutatava kaardi nii üldist kui ruumilist ja klassisisest klassifitseerimisõigsust. On leitud, et suurema klassifitseerimistäpsusega

piltidelt arvutatud maastikuindeksite varieeruvuskoefitsient on väiksem (Shao *et al.* 2001; Altamirano *et al.* 2012). Teisalt, kuigi üldine klassifitseerimistäpsus võib olla hea, ei ütle see midagi vigade ruumilise või klassidevahelise jaotumise kohta, mistõttu klassifitseerimisvigadest tingitud ebamäärasust indeksite väärtustele on raske ennustada (Shao *et al.* 2001; Fassnacht *et al.* 2006). Klassifitseerimisvigadest tekkivat ebamäärasust hinnatakse tavaliselt veamaatriksi abil, mis kirjeldab iga maakatte klassi klassifitseerimise täpsust ja iga klasside paari suhtelist valesti klassifitseerimist. Kuna veahinnangud on kulukad, pole paljudele kaugseire andmetel põhinevatele kaartidele neid üldse tehtud, mis kahandab ka nendelt arvutatud indeksite usaldusväärsust (Shao, Wu 2008). Hinnangud klassifitseerimise täpsusele peaksid siiski sisalduma igas kaugseire andmestikku kasutavas uurimuses. Klassifitseerimistäpsuse parandamiseks tasub kasutada objekt-orienteeritud, homogeensetel objektidel põhinevaid klassifitseerimisalgoritme. Ka klasside vähendamine / kokku panemine võib aidata saavutada paremat täpsust (Fassnacht *et al.* 2006).

On leitud, et klassifitseerimistäpsuse suhtes on tundlikumad klassi taseme indeksid maastiku taseme omadega võrreldes ning eraldiste struktuuri kirjeldavad indeksid kompositsiooni kirjeldavate indeksitega võrreldes (Mas *et al.* 2010; Altamirano *et al.* 2012). Ka Lechner *et al.* (2013) ja Langford *et al.* (2006) jõudsid järeldusele, et erinevused ruumilises mitmekesisuses ja autokorrelatsioonis mõjutavad indeksite väärtusi palju rohkem kui klasside suhtelise osakaalu ehk kompositsiooni muutmine maastikus.

2. Andmed ja metoodika

2.1. Satelliidipildid

Töös kasutati kahte multispektraalset satelliidipilti: RapidEye (pildistatud 3.07.2010, kell 10:27 GMT) ja GeoEye (pildistatud 7.11.2010, kell 09:37 GMT). RapidEye toodab suure ruumilise lahutusega multispektraalseid satelliidipilte, mille ortorektifitseeritud pildi piksli suurus on 5 m. RapidEye satelliitide sensoritel on 5 spektraalkanalit: sinine (440-510 nm), roheline (520-590 nm), punane (630-685 nm), ülemineku kanal punaselt lähisinfrapunasele (*Red Edge*, 690-730 nm) ja lähisinfrapunane kanal (760-850 nm) (RapidEye Standard Image Product Specifications 2013). Sensoril on 12-bit dünaamiline vahemik, mis annab piksli heledusväärtuste (ingl *Digital Number*, DN) vahemikuks 4096. Piltidele on tootja poolt tehtud radiomeetriline korrektsioon ja heledusväärtused on omakorda teisendatud 16-bit vahemikku, muutes suhtelised DN väärtused absoluutseteks kiirgusväärtusteks. Atmosfääri korrektsioone ei ole piltidele tehtud.

GeoEye (alates 2013. a jaanuarist kuulub DigitalGlobe'i alla) toodab suure ruumilise lahutusega multispektraalseid ja pankromaatilisi satelliidipilte, mille ortorektifitseeritud piltide ruumiline lahutus on vastavalt 2 m ja 0.5 m. Sensoril on 5 kanalit: pankromaatiline (450-800 nm), sinine (450-510 nm), roheline (510-580 nm), punane (655-690 nm) ja lähisinfrapunane (780-920 nm). Käesolevas töös kasutati multispektraalse kanali pilti. Pikslite heledusväärtuste vahemik on 8-bit ehk 256.

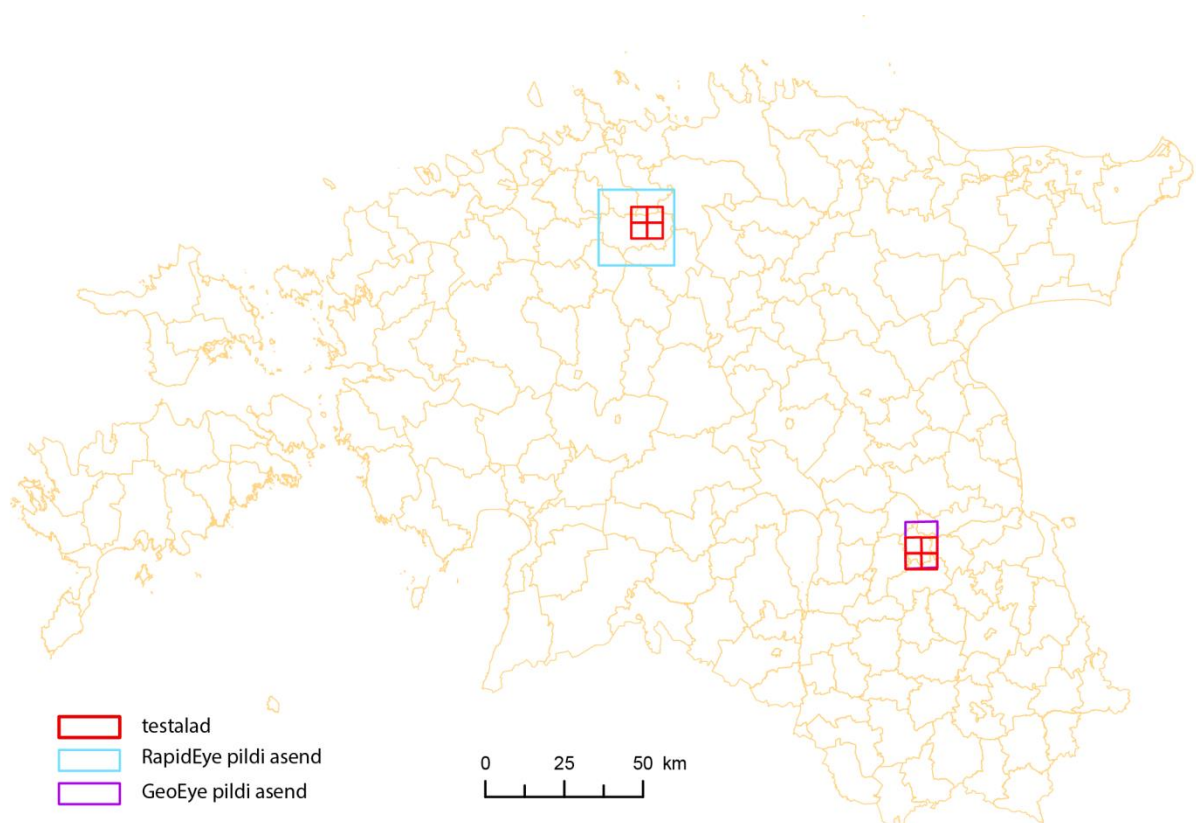
Mõlemad pildid on tootja poolt georeferentseeritud Mercatori universaalsesse põikprojektsiooni, tsooni 35. Piltidelt lõigati välja kaks testala suurusega 10x10 km, mis pärast tarkvaratöötlust jaotati indeksite arvutamise tarvis omakorda neljaks väiksemaks alaks, kokku seega kaheksaks alaks. RapidEye pildi testala kujutab Kose ümbrust Harjumaal (joonis 1, joonis 3), GeoEye testala asub Tartust lõunas (joonis 2, joonis 3). Testalad lõigati välja programmi Erdas Imagine tööriistaga Subset, määrates vähimad ja suurimad X,Y koordinaadid.



Joonis 1. Kose ümbrust kujutav testala RapidEye satelliidipildil



Joonis 2. Tartu linnast lõunasse jääv testala GeoEye satelliidipildil



Joonis 3. Satelliidipiltide ja neilt välja lõigatud testalade asend.

2.2. Tarkvarad

Satelliidipiltide klassifitseerimiseks kasutati järgmisi tarkvarasid: Erdas Imagine 2011, ENVI 5, Idrisi Selva, eCognition. Erdas Imagine on ettevõtte Intergraph loodud kaugseire pilditötlustarkvara rasterandmete töötlemiseks ja analüüsimiseks. ENVI on samuti pilditötlustarkvara, mida peamiselt kasutavad kaugseire spetsialistid, arendajaks on Exelis VIS. Idrisi on Clark Labs poolt arendatud GIS-i ja kaugseire integreeritud tarkvara, mida kasutatakse nii GIS ruumianalüüsiks kui pilditötluseks. eCognition on ettevõtte Trimble objektorienteeritud pilditötlustarkvara, mis kasutab objektide ehk pikslite gruppide põhiseid algortime erinevalt traditsioonilistest pikselhaaval töötavatest meetoditest. Pikslid ei vaadata mitte eraldi, vaid arvestatakse nende konteksti.

Maastikuindeksite arvutamiseks kasutati Kevin McGarigali ja Barbara Marksi loodud programmi Fragstats, mis on üks populaarsemaid maastiku struktuurianalüüsise tegemisel (McGarigal, Marks 1995). Fragstats pakub laialdast maastikuindeksite valikut ja sellest on kaks versiooni: vektor- ja rasterkujul andmete tarvis. Käesolevas töös kasutati rasterandmete versiooni, mis töötab erinevate rasterformaatidega, sh ASCII, ESRI grid / raster, Erdas Imagine grid, GeoTIFF grid. Nii raster- kui vektorversioonis on võimalik arvutada samu indekseid, kuid rasterversioon võimaldab veel mitmeid lähima naabri indeksite arvutusi (McGarigal, Marks 1995).

2.3. Etalonideta klassifitseerimine

Etalonideta klassifitseerimismeetodit (*unsupervised classification*) ehk näidisteta klassifitseerimist või klasterdamist saab kasutada maakatte kaardistamiseks ja see põhineb arvutitarkvara, täpsemalt algoritmide võimel tunda ära sarnaseid alasid. Erinevalt näidistega klassifitseerimisest (*supervised classification*), mis kasutab õpetuspiirkondi ja juba teadaolevaid maakatteklasse, jagatakse näidisteta klassifitseerimisel sarnaste spektraalväärtustega heledusväärtuste teljel tihedamalt koos paiknevad pikslid klastritesse, millele kasutaja annab hiljem sisulise tähenduse. Klasterdamist kasutatakse peamiselt andmekaeveks, mustrite uurimiseks, andmete rühmitamiseks (Jain *et al.* 1999). Näidisteta klassifitseerimisel on oluline roll interpreteerijal, kes peab algoritmi loodud klastritele nimed või klassid määrama. Selle klassifitseerimismeetodi puuduseks on pikslite ruumilise ühtluse ja paiknemisega mitte arvestamine.

Populaarsed klasterdamise algoritmid on ISODATA ja k-means, kus pikslid jagatakse pikselhaaval klastritesse lähima spektraalse kauguse (eukleidilise kauguse) alusel. Iga iteratsiooniga arvutatakse klastrite keskmised spektraalväärtused ümber ja vastavalt sellele võivad pikslid muuta oma klasse. Klassifitseerimisalgoritm lõpetab töö, kui maksimaalne määratud iteratsioonide arv on täidetud või kui maksimaalne määratud osakaal pikslid säilib iteratsiooni käigus muutumatuna, uude klastrisse määramata. Erinevalt k-means algoritmist lubab ISODATA määrata klastrite täpse arvu asemel võimaliku vahemiku. Seetõttu saavad klastrid iteratsioonide käigus jaguneda ja liituda. Klastrite jagunemine toimub, kui standardhälve on lubatust suurem, klastris on jagunemiseks piisav arv pikslid ning teatud hulga pikslite heledusväärtused erinevad piisavalt palju klatri keskvaartusest (Ball, Hall 1965).

K-means on üks enim kasutatavatest klasterdamise algoritmidest, mida on lihtne kasutada ja mis ei ole eriti keerukas – algoritmi keerukus on $O(n)$, kus n on pikslite arv (Jain *et al.* 1999). Algoritmil on kolm parameetrit: klastrite arv K , klastrite algatamise

reegel ning kauguse arvutamise viis klasteri heledusväärtuste keskpunkti ja teiste punktide vahel (Jain 2010). Kõige olulisem on neist K valik, millele sobiva väärtuse leidmine nõuab enamasti katsetamist. Erinevad klasterite algatamise reeglid võivad viia eri lõpptulemusteni, sest k-means läheneb vaid lokaalse miinimumini. Ka siin võib katsetada erinevate algatamise reeglitega ja valida kõige väiksema ruutveaga jagunemine. Kauguse arvutamiseks heleduväärtuste vahel on levinuim Eukleidilise kauguse kasutamine, kuid võimalikud on ka teised viisid, näiteks Mahalanobise kaugus. Keskvaertusele lähima kauguse alusel hakatakse pikslid klasterites ümber jaotama, kuni saavutatakse lävend, kus pikslid ei muuda iteratsiooni käigus oma klassi või ruutviga ei vähene enam märgatavalt. ISODATA (*Iterative Self-Organizing Data Analysis*) algoritm on k-means algoritmi modifikatsioon (Jain 2010). Põhiline erinevus on selles, et ISODATA algoritm võimaldab jätta klasterite arvu täpselt defineerimata, anda vahemiku.

Erdas Imagine 2011 tarkvarapakettis on näidisteta klassifitseerimise tarvis üks algoritm – ISODATA, ENVI 5 tarkvara pakub lisaks sellele ka k-means algoritmi, Idrisi Selva pakettis on neid mitu: ISODATA, k-means, cluster, isoclust, maxset, chaincluster. Kuigi samanimelised, võivad eri tarkvarades nende algoritmide võimalikud parameetrid erineda (tabel 1).

Tabel 1. Algoritmi ISODATA parameetrid Erdas Imagine, ENVI ja Idrisi tarkvarapakettides.

	Erdas Imagine 2011	ENVI 5	Idrisi Selva
ISODATA			
Soovitud klasterite / klasside arv	x	x	x
Maksimaalne iteratsioonide arv	x	x	x
Klassi mitte muutvate pikslite % lävi	x	x	x
Klasterite keskpunktide algatamise reegel	x		x
Klasteri minimaalne suurus		x	x
Suurim lubatud standardhälve klassis	x	x	x
Vähim lubatud eukleidiline kaugus klasterite vahel		x	x
Suurim iteratsiooni käigus ühendatavate paaride hulk klassis		x	x
Suurim lubatud standardhälve keskmisest (valikuline)		x	
Suurim lubatud kauguse viga (valikuline)		x	

ISODATA algoritmi katsetati kolmes programmis, püüdes anda samaväärsed parameetriväärtused (tabel 2). Klasterite liitmise ja lahutamisega seotud parameetriväärtusteks võeti vaikimisi pakutavad. Erdas Imagine 2011 tarkvaras saab erinevalt ENVI ja Idrisi ISODATA moodulitest määrata loodavate klasterite arvuks vaid ühe väärtuse, mitte vahemiku, mistõttu on see ISODATA algoritm pigem k-means algoritmi sarnane. Sellest tulenevalt määrati kõikide tarkvarade puhul klasterite võimalikuks arvuks üks väärtus. Kuna kasutaja ei saa defineerida loodavate klasterite vahemikku, siis on Erdas ISODATA võimalused piiratud, kuna klasterite liitmise ja lahutamisega seotud parameetreid ei ole tarvis. Kose jaoks valiti klasterite arvuks 6, kuid

Tartu puhul sobis klastrite arv 5 paremini: klassifitseerimistulemused eri programmides ühtisid paremini ja tulemused olid kergemini interpreteeritavad.

Tabel 2. Testalade klassifitseerimise parameetrid algoritmiga ISODATA Erdas Imagine, ENVI ja Idrisi Selva tarkvarades.

ISODATA	Erdas 1	Erdas 2	ENVI	Idrisi 1	Idrisi 2	Idrisi 3
Klastrite arv	5, 6*	5, 6*	5-5, 6-6*	5-5, 6-6*	5-5, 6-6*	6-6
Max iteratsioonide arv	8	8	8	8	8	8
Klassi mitte muutvate pikslite %	95	95	95	95	95	95
Klastrite keskpunktide alagatamise reegel	diagonal axis	principal axis	-	diagonal axis	random partition	random seeds
Klastri min suurus	-	-	1	1	1	1
Max standardhälve klassis	1	1	1	1	1	1
Min Eukleidiline kaugus klastrite vahel	-	-	5	5	5	5
Max ühendatavate paaride hulk iteratsiooni käigus	-	-	2	2	2	2

* Kose – 6 klastrit; Tartu – 5 klastrit.

Erinevalt tarkvarast ENVI pakuvad Erdas 2011 ja Idrisi Selva klastrite keskpunktide alagatamise reegleid. Idrisi-s on nii ISODATA kui k-means'i jaoks kolm klastrite keskpunktide alagatamise reeglit: **juhuslik jaotus** (ingl *random partition*; pikslid määratakse juhuslikult klastritesse, mille alusel arvutatakse klastrite keskpunktid), **juhuslikud keskpunktid** (ingl *random seed*; esmalt määratakse klastritele juhuslikud keskpunktid ja keskpunktidele lähima kauguse alusel jaotatakse pikslid klastritesse), **diagonaalne telg** (ingl *diagonal axis*; klastrite keskpunktid määratakse ühtlase jaotusega heledusväärtuste teljel n-mõõtmelises ruumis, kus n on spektraalkanalite arv). Erdas tarkvaras saab valida vaikimisi pakutava diagonaalse telje ja **peatelje** (ingl *principal axis*; keskpunktid arvutatakse ühtlase jaotusega esimese peakomponendi teljel) vahel. Teoorias ei tohiks lõpptulemus esialgsete keskpunktide valikust eriti sõltuda, kuid see võib aidata märgatavalt vähendada iteratsioonide arvu, jõudmaks sobiva tulemuseni (Ball, Hall 1965). Erinevad autorid on aga tõestanud klassifitseerimistulemuste sõltuvust algsetest klastrite heleduskeskväärtustest ja pikslite jagunemisest klastrite vahel (Peña *et al.* 1999; Huang 2002; Redmond, Heneghan 2007).

Kuna Idrisi ja ENVI ISODATA moodulid lubavad määrata klastrite arvu vahemikku, siis võrreldi neid eraldi nii, et määrati klastrite vahemikuks 1-6, teistele parameetritele määrati tabelis 2 märgitud väärtused.

K-means mooduleid saab võrrelda Idrisi ja ENVI tarkvarapakettides (tabel 3). Kuna Erdas ISODATA moodul paistab olevat samalaadne k-means algoritmiga, siis võrreldi ka seda eelmainitutega. K-means eeldab üldiselt, et klastrite arv on teada, mistõttu puuduvad sellel mitmed klastrite liitmise ja lahutamise seotud parameetrid, muus osas kattuvad parameetrid tarkvarades ISODATA parameetritega. Idrisi pakettis ei saa siiski määrata klastrite kindlat arvu, vaid maksimaalset klastrite arvu, mistõttu loodavate klasside arv ei ole täpselt teada – algoritm ise määrab selle. Lisaks pakub Idrisi ka siin klastrite keskpunktide alagatamise reeglit ning ENVI lubab soovi korral

määrata suurima lubatud standardhälbe keskmisest ning suurima lubatud kauguse vea, mis jäeti vaikimisi tühjaks. Kuna Idrisi diagonaalse telje klastrite algatamise reeglina klassifitseeris algoritm mõlemal alal pikslid vaid kolme klassi, siis seda algoritmide töö võrdluses ei kasutatud.

Tabel 3. Klassifitseerimise parameetrid algoritmiga k-means kolmes tarkvarapaketis.

K-means	Erdas (ISODATA)	ENVI	Idrisi 1	Idrisi 2
Klastrite arv*	6	6	6	6
Max iteratsioonide arv	8	8	8	8
Klassi mitte muutvate pikslite %	95	95	95	95
Klastrite keskpunktide algatamise reegel	diagonal axis	-	random seeds	random partition

*Idrisi paketi maksimaalne klastrite arv

2.4. Etalonidega klassifitseerimine

Näidistega klassifitseerimisel on oluline luua esmalt õpetuspiirkonnad, mille alusel hakatakse pikseleid klassidesse jaotama. Kasutaja peab valima iga klassi jaoks esinduslikud pikslid ehk näidisalad, mis aitaksid vastava klassi mustreid ära tunda. Selleks saab kasutada ka lisaandmeid, nagu kaardid, aerofotod, maapealsed mõõdistused. Klasside arv peab olema eelnevalt teada. Õpetuspiirkondadega luuakse signatuurid, iga signatuur vastab ühele klassile. Signatuurid on kas parameetrilised või mitteparameetrilised. Esimese põhineb pikslite statistilistel parameetritel (näidisalade põhjal luuakse keskmiste heledusväärtuste vektorid, hajuvusmaatriks) ja iga piksel määratakse lähimasse spektraalsesse klassi, teine on sõltumatu andmete omadustest (ei põhine statistikal) ja piksel määratakse klassi vaid juhul, kui tema heledusväärtus jääb signatuuri piiridesse. Mitteparameetrilised signatuurid aitavad parandada klassifitseerimisetäpsust normaaljaotusele mitte vastavate klasside ja keerukate maastike puhul (Lu, Weng 2007). Õpetuspiirkondade valikul on oluline minimeerida spektraalset ülekattet eri klasside vahel (õpetuspiirkonnad peaksid olema seesmiselt homogeensed) ning samal ajal proovida kaasata andmestiku kogu spektraalset gradienti (Ediriwickrema, Khorram 1997). Klassifitseerimine toimub signatuuride alusel, kasutades sobivat klassifitseerimise otsustusreeglistikku ehk matemaatilist algoritmi, mis jaotab pikslid eri klassidesse.

Käesolevas töös kasutati näidistega klassifitseerimiseks lühima tee (ingl *Minimum Distance*) ja suurima tõepära (ingl *Maximum Likelihood*) algoritme ja võrreldi nende tööd kolmes tarkvaras: Erdas Imagine, ENVI, Idrisi Selva. Suurima tõepära klassifikaator on üks enim kasutatud klassifitseerimisalgoritme (Lu, Weng 2007) ja lühima tee klassifikaator sarnaneb valikus olnud näidistega klassifitseerimise algoritmidest kõige enam ISODATA näidisteta klassifitseerimise algoritmile. Lühima tee klassifikaator arvutab spektraalse kauguse (Eukleidilise kauguse) klassifitseeritava piksli ja iga signatuuri keskmise heledusvektori vahel. Suurima tõepära klassifikaator eeldab spektraalkanalite andmete histogrammide normaaljaotust. Põhineb tõenäosusel, et piksel kuulub konkreetse klassi. Vaikimisi on need tõenäosused kõigi klasside jaoks võrdsed. Algoritm arvestab õpetuspiirkondade signatuuride keskmisi väärtusi ja erinevalt lühima tee klassifikaatorist ka hajuvust.

Näidistega klassifitseerimise tarvis valiti Erdas ISODATA mooduliga näidisteta klassifitseerimise tulemusena saadud mõlema testala pildilt õpetuspiirkonnad, iga klassi jaoks üks, mis iseloomustaksid loodavate klasside spektraalseid väärtusi, ja digitaliseeriti need vektorfaili. Nii loodi parameetrilised signatuurid. Erdas signatuuride toimetaja lubab enne klassifitseerimist koostada veamaatriksi, kus on õigesti klassifitseeritud näidisalade pikslite osakaalud mõnda Erdas-e pakutavat klassifitseerimisalgoritmi kasutades. Kose pildi jaoks olid need lähima kauguse algoritmi kasutades 97,2-100% ja maksimaalse tõepära algoritmi kasutades 98,66-100%, Tartu jaoks vastavalt 96,8-99,99% ning 99,88-100%. Eri klasside signatuuride eristumist üksteisest võrreldi eri spektraalkanalites, kasutades signatuuride võrdlemise diagramme ja kauguste erinevuste maatriksit (*Transformed Divergence*). Üldiselt eristusid signatuurid üksteisest väga hästi, eriti lähisinfrapunases kanalis ja Kose puhul ka ülemineku kanalis punaselt lähisinfrapunasele.

Tarkvaraga Erdas loodud signatuurid toodi Idrisi tarkvarasse üle tööriista MAKESIG abil. MAKESIG küsib ka minimaalset näidisala suurust klassi kohta pikslites, vaikimisi on see 10 x spektraalkanalite arv. Käesolevas töös kasutatud näidisalade suurus oli kõigi klasside jaoks suurem. ENVI jaoks oli vaja iga signatuur salvestada eraldi .shp vektorformaati, mille järel sai need importida ENVI vektorformaati .evf ja kasutada õpetuspiirkonnadena (ENVI-s *Region of Interest* ehk ROI).

Erdas-e näidistega klassifitseerimiseaknas saab valida klassifitseerimisalgoritmide vahel (parameetrilised ja mitteparameetrilised) ning suurima tõepära klassifikaatori valimisel teha linnuke ka klassi tõenäosuste kasutamise kasti. ENVI lubab lähima kauguse klassifitseerimiseaknas soovi korral määrata ka suurima lubatud standardhälbe keskmisest ning suurima lubatud kauguse vea, mõlemad iga klassi kohta eraldi või kõikidele klassidele sama. Idrisi lubab defineerida suurima otsingukauguse heledusväärtuste DN ühikutes. Nende parameetrite määramisel ei pruugi kõik pikslid saada klassifitseeritud. Ka ENVI ja Idrisi suurima tõepära klassifikaatori parameetrite aknas saab määrata klassidele tõenäosused. Käesolevas töös on kõigi klasside tõenäosused jäetud vaikimisi võrdsed ning lühima tee klassifikaatori lubatud kaugusele ei ole seatud piiranguid.

2.5. Objektorienteeritud klassifitseerimine

Objektorienteeritud klassifitseerimiseks kasutati selleks spetsiaalselt loodud tarkvara eCognition ning Idrisi Selva moodulit *Segmentation Classifiers*. Mõlemas tarkvaras tuleb enne klassifitseerimist pilt segmenteerida ehk osadeks jagada, mis on edasise klassifitseerimise eelduseks. Segmenteerimisel ühendatakse sarnaste omadustega naaberpikslid. Selleks pakub eCognition erinevaid algoritme. Osadeks jaotamisel arvestatakse pikslite radiomeetriat ning naaberalade homogeensust, mis sõltub kujust ja kompaktsusest. Segmenteerimisel tarkvaraga eCognition kasutati pikslipõhist segmenteerimist ja algoritmi *Multiresolution Segmentation*, mis põhineb pikslite radiomeetrial ning kuju sarnasusel (eCognition Developer 8.9 User Guide 2013). Kasutaja peab määrama mõõtkava, kuju ja kompaktsuse parameetrid. Kui mõõtkava parameetri väärtus on madal, peavad objektid olema seesmiselt väga ühetaolised, suurem väärtus lubab suuremat heterogeensust. Kuju ja kompaktsuse parameetritele saab anda väärtusi 0-0,9, mis näitab vastava parameetri suhtelist kaalu, võrreldes spektraalse väärtusega. Kuna RapidEye ja GeoEye on erineva ruumilise lahutusega, siis määrati Kose ja Tartu piltide objektorienteeritud klassifitseerimiseks erinevad mõõtkava parameetrid (Kose 300, Tartu 30), kuju ja kompaktsuse parameetrid väärtustati 0-ga, kuna Idrisi Selvas neid ei ole. Idrisi segmenteerimistööriist

SEGMENTATION töötab pikslite radiomeetriapõhiselt. Loodavate segmentide suuruse saab määrata parameetriga sarnasuse lubatud hälve (*similarity tolerance*).

Näidistega klassifitseerimiseks on eCognitionis lähima naabri mitteparameetriline klassifitseerimise meetod (*Nearest Neighbor Classification*). Esmalt peab iga klassi jaoks valima segmenteeritud osade seast esinduslikud näidisalad, mille alusel algoritm määrab ülejäänud osad klassidesse. Kasutaja saab määrata objektide karakteristikuruumi (*feature space*). Käesolevas töös määrati klassifitseerimisel arvestatavateks karakteristikuruumi parameetriteks keskmine heldusväärtus ja standardhälve.

Idrisi Selvas püüti valida samad näidisalad, mis tarkvaras eCognition, kuid alade piirid võisid veidi erineda, sõltuvalt tekitatud segmentidest. Objektide klassifitseerimiseks on Idrisi-s lisaks õpetuspiirkondadele tarvis ka juba klassifitseeritud etalonpilti, mis loodi samade õpetuspiirkondade põhjal klassifikaatoriga KNN (*k-nearest neighbors*), kus parameeter k ehk lähimate naabrite arv väärtustati 1-ga. Sellega peaks olema algoritm KNN samaväärne lähima naabri klassifitseerimise meetodiga, mida kasutati ka eCognitionis. Etalonpildi alusel klassifitseeritakse segmendid enamuse reegli (ingl *majority rule*) põhjal kasutaja määratud klassidesse.

2.6. Maastikuindeksite arvutamine tarkvaraga Fragstats

Programmis Fragstats saab maastikuindekseid arvutada kolmel tasemel: eraldise, klassi ja maastiku tasemel. Käesolevas töös arvutati indeksid maastiku tasemel, kuna eraldise ja klassi taseme arvutused ja analüüs oleksid olnud väga aeganõudvad ja mahukad suure ruumilise lahutusega piltide nii paljude eri tulemuste jaoks. Väärtused arvutati pea kõigile indeksitele (111-le) (lisa 1), et oleks võimalik võrrelda kõikide indeksite varieeruvust ning ka seda, kas indeksid, mis kirjanduse järgi korreleeruvad, varieeruvad sarnaselt. Erinevad autorid on küll püüdnud välja selgitada indeksite tuumikut, mis kirjeldaksid maastikumustri erinevaid aspekte, kuid enamasti ei ole kaasatud kõiki indekseid ning konsensust pole saavutatud (Lechner *et al.* 2013). Käesolevas töös jäeti välja siiski kogu pindala indeks (TA) ja eritüübiliste eraldiste rohkust kirjeldavad mitmekesisuse indeksid (PR – eritüübiliste eraldiste rohkus, PRD – eritüübiliste eraldiste tihedus, RPR – suhteline eritüübiliste eraldiste rohkus), kuna need eraldised põhinevad pindalal ja eritüübiliste eraldiste arvul, mis käesolevas töös olid võrreldavatel aladel samad, kasutaja poolt kindlaks määratud. Neid väärtusi ei muutnud ka tarkvaratöötlus. Seetõttu oli nende indeksite variatsioonikoefitsient kõikidel juhtudel 0. Samasugustele järeldustele jõudsid ka Mas *et al.* (2010).

Funktsionaalsete indeksite arvutamiseks on tarvis täpsustavaid parameetreid, mis üldiselt sõltuvad uuritavast nähtusest ja mõjutavad indeksi väärtust. Kontrastsuse ja sarnasuse indeksite jaoks peab kasutaja defineerima sarnasuse ja serva kontrastsuse maatriksid, mis kirjeldavad, kui erinevad või lähedased on eraldiste klassid maakatte/-kasutuse poolest omavahel. Selleks tuleb klasside paaridele anda kaalud vahemikus 0 kuni 1, kus 0 tähendab, et kontrastsus või sarnasus puudub ja 1, et need on maksimaalsed (lisa 2). Lähedus-, sarnasus- ja sidususindeksi tarvis on vaja määrata otsimisraadius ehk lävendkaugus, mille raadiuses arvutused tehakse. Kose klassifitseeritud piltide jaoks määrati nende väärtuseks 500 m, Tartu jaoks 200 m. Tuumala indeksite jaoks peab kasutaja defineerima servasügavuse. See võib olla kõikide servade jaoks sama, aga võib ka erineda, sõltuvalt milliste maakattetüüpide servad koos paiknevad. Käesolevas töös määrati fikseeritud servasügavus 100 m Kose alal ja 50 m Tartu alal.

2.7. Klassifikatsioonide ja indeksite võrdlemine

Eri tarkvarades või eri viisil klassifitseeritud Kose ja Tartu pilte võrreldi veamaatriksite abil, mis võimaldas võrrelda nii üldist klassifitseerimistäpsust kui hinnata vigade jaotumist klasside lõikes. Võrreldi näidisteta, näidistega ning objektorienteeritud klassifitseerimisi eri tarkvarades. Võrdlemaks eri viisil või erinevates programmides klassifitseeritud satelliidipiltidelt arvutatud indeksite tundlikkust, kasutati variatsioonikoefitsente. Selleks arvutati variatsioonikoefitsendid Kose, Tartu ja nende veerandite pildipaaride indeksitele.

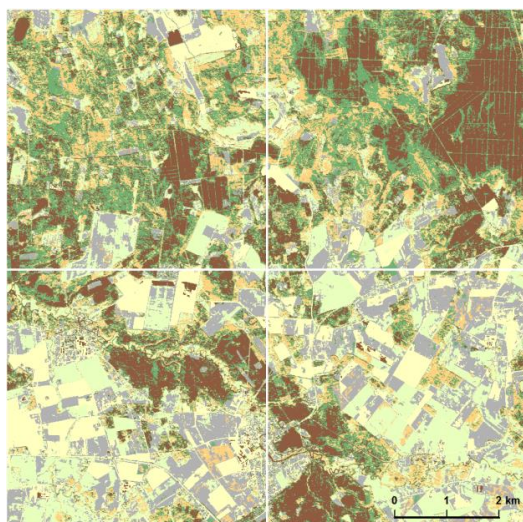
3. Tulemused

3.1. Etalonideta klassifitseerimine

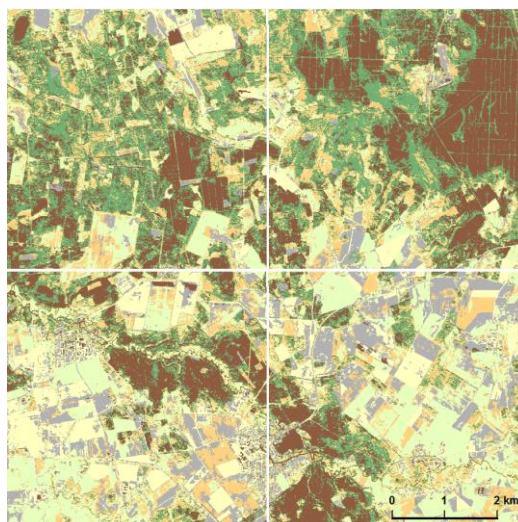
Mõlema testala klassifitseerimisel ISODATA algoritmiga olid Erdas Imagine diagonaalse telje klasterite algamise reeglit kasutades saadud tulemused ja ENVI tulemused üksteisega vastavuses, Tartu puhul ühildusid tulemused 100%, Kose puhul 99,97% (tabel 4). Idrisi Selva samanimeline algoritm aga andis mõnevõrra teistsugused tulemused. Kasutades klasterite keskpunktide algamise reegleid *diagonal axis* ja *random partition*, kõikus üldine ühildumine Erdas (aga ka ENVI) tulemustega kahe ala peale kokku 62-98%, kusjuures pea iga kord olid klassifitseerimise tulemused ettearvamatud ja erinevad (tabel 4). Kui juhuslike keskpunktide algamise reeglit (*random seed*) kasutati, siis ühtisid Kose klassifitseerimise tulemused üsna hästi – 90,3% (korduval katsetamisel). See reegel ei sobinud aga Tartu pildi puhul teiste tarkvaradega võrdlemiseks, sest kuigi tekkis soovitud 5 klassi, oli kolmes klassis kokku alla 0,1% pildi kõikidest pikslitest. Kasutades klasterite algamiseks põhikomponentide telge tarkvaraga Erdas, erinesid sama tarkvara diagonaalse telje reeglit kasutades tulemused 18% Kose puhul ja vaid 2% Tartu puhul. Kõikide tulemuste puhul katsetati algoritmide tööd korduvalt. Maastikuindeksite varieeruvuse arvutamiseks valiti välja kolm tarkvarades Erdas ja Idrisi algoritmiga ISODATA klassifitseeritud Kose ala kujutavat pilti, kasutades erinevaid klasterite algamise reegleid, mis erinesid üksteisest piisavalt, kuid olid saadud algoritmi korduval katsetamisel samade parameetritega (joonis 4).

Tabel 4. Kose ja Tartu testalade ISODATA moodulitega loodud klassifikatsioonide ühilduvus eri tarkvarades.

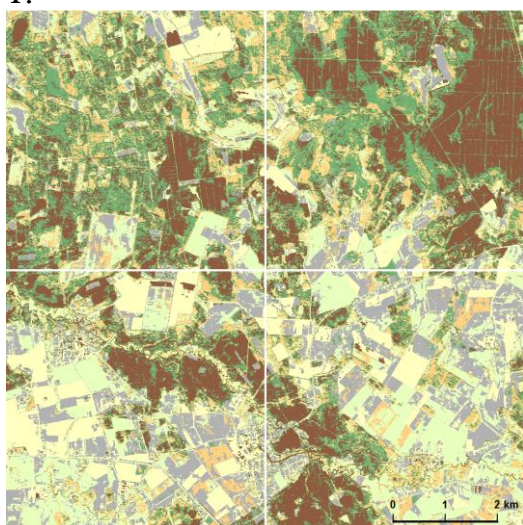
KOSE	ENVI	Erdas 2 (<i>princ. axis</i>)	Idrisi 1 (<i>diag. axis</i>)	Idrisi 2 (<i>rand. part.</i>)	Idrisi 3 (<i>rand. seed</i>)
Erdas 1 (<i>diag. axis</i>)	99.97%	82%	62, 68, 74, 90%	65, 80, 89, 94%	90.3%
TARTU	ENVI	Erdas 2 (<i>princ. axis</i>)	Idrisi 1 (<i>diag. axis</i>)	Idrisi 2 (<i>rand. part.</i>)	
Erdas1 (<i>diag. axis</i>)	100%	98%	68, 77, 81, 82, 90, 92, 93%	82, 85, 90, 93, 94, 95, 98%	



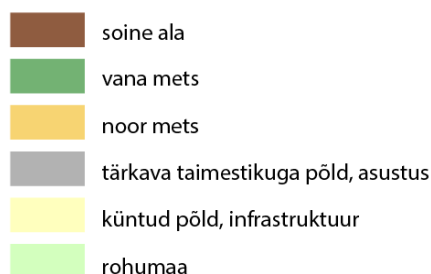
1.



2.



3.



Joonis 4. ISODATA algoritmiga klassifitseeritud Kose testala tarkvarades **Erdas** (*diagonal axis*) (1), (*principal axis*) (2) ja **Idrisi** (*random seed*) (3).

Idrisi ja ENVI ISODATA moodulite võrdlemisel klastrite vahemiku määramisega 1-6, klassifitseeris ENVI algoritm nii Kose kui Tartu jaoks 3 klassi, Idrisi algoritm kõikide keskpunktide algatamise reeglite ja samade parameetriväärtuste korral 6 klassi, mistõttu neid klassifitseerimisi ei võrreldud.

K-means moodulitega klassifitseerimise tulemustest selgub, et ENVI k-means vastab täielikult Erdas ISODATA algoritmile, mis on ka ootuspärane, kuna nende tarkvarade ISODATA algoritmide ühtisid niisamuti (tabel 5). ISODATA algoritmiga klassifitseerides ei antud klastrite arvu defineerimisel vahemikku, vaid määrati üks arv, millega võrdsustati nende algoritmide töö. Seetõttu on ka ENVI ISODATA ja k-means moodulite tulemused samaväärsed üheselt määratud klastrite arvu korral. Nii Tartu kui Kose alal ühtisid need ENVI kaks näidisteta klassifitseerimise algoritmi 100%. Idrisi k-means mooduliga klassifitseerides tulid erinevad tulemused, sõltuvalt klastrite algatamise reeglist. Kasutades juhuslike keskpunktide reeglit (*random seed*), ühtisid mõlemal testala tulemused ENVI omadega 76%, kuid Tartu alal lõi algoritm vaid 5 klastrit, kuna Idrisi-s peab määrama suurima klastrite arvu (algoritm võib neid ka vähem tekitada). Juhusliku jaotuse (*random partition*) reegluga klassifitseeris algoritm aga mõlemal alal pikslid kord viide, kord kuude klassi, kuid kummagi arvu korral, alati

ühtemoodi. Seetõttu klassifitseeriti ka ENVI k-means algortimiga mõlemad alad viide klassi, et saaks võrrelda tulemusi nii viie kui kuue klassi korral. 5 klassi korral ühtisid kahe tarkvara tulemused paremini. Kuna 6 klassi korral ei olnud Tartu alal klastrite 3 ja 4 võrdlemisel sisulist vahet – mõlemad esindasid sarnase maakattega alasid (taimkatteta põllud), siis liideti need klassid kokku ja saadi 5 klastrit. Üldine ja klassidevaheline klassifitseerimistäpsus paranes sellest. Üldiseks klassifitseerimistäpsuseks saadi Tartu jaoks **81%**. Klasside lõikes erinesid kahe tarkvara tulemused kohati üsna palju, isegi üle 50% (tabel 6-7). Mõlemal alal põhjustas erinevusi taimkattega põldude ja rohumaade klasside spektraalne sarnasus. Tartu alal parandas klassifitseerimise ühtivust klasside liitmine. Diagonaalse telje reeglit (*diagonal axis*) kasutades tekitas algoritm vaid kaks klassi ja seda tulemust ei kasutatud võrdluses.

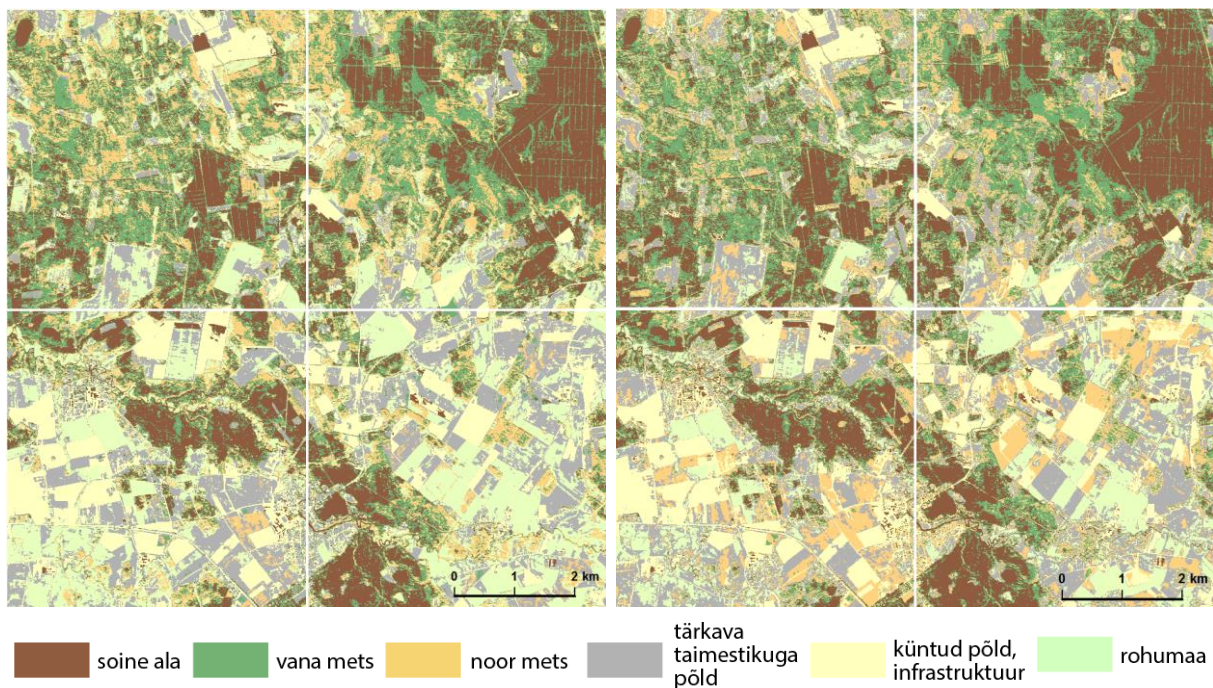
Kuna GeoEye pildi ruumiline lahutus on väga suur (2 m), siis tekkis etalonideta klassifitseerimisel palju üksikuid mürapiksleid (ingl *salt and pepper effect*), mille spektraalne väärtus erines ümbritsevatest pikslitest. Mürapikslite eemaldamiseks sobib hästi mediaanfilter. Tartu ENVI ja Idrisi k-means algoritmidega klassifitseeritud piltidele rakendati 5x5 mediaanfiltrit, eemaldamaks mürapiksleid ja maastikuindeksite efektiivsemaks arvutamiseks. Mürapikslite korral oleks eraldisi väga palju tulnud ja arvutused väga aeglaseks muutunud. Klassifitseerimistulemuste üldist ja klassidevahelist ühtivust filtri kasutamine peaaegu ei muutnud (üldine jäi samaks, klasside ühtivus muutus kuni 3%).

Tabel 5. Kose ja Tartu testalade k-means moodulitega loodud klassifikatsioonide ühilduvus eri tarkvarades

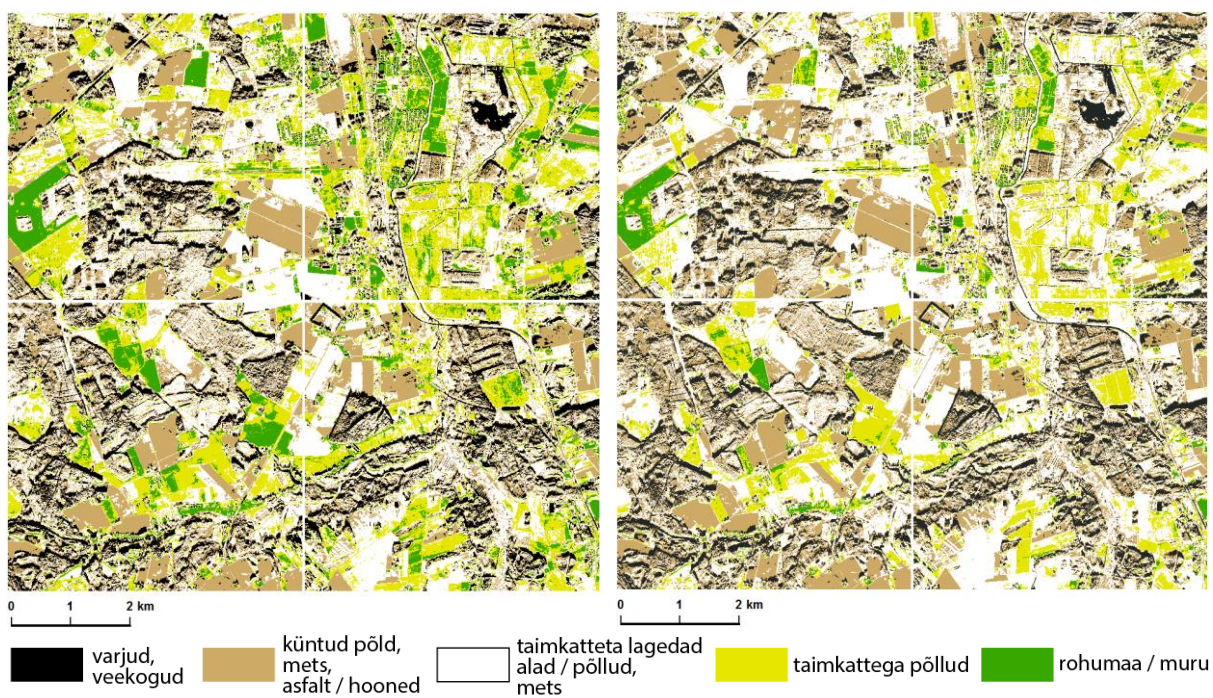
KOSE	Erdas (<i>ISODATA</i>)	Idrisi 1 (<i>rand. seed</i>)	Idrisi 2 (<i>rand. part.</i>)
ENVI	99.96%	76%(6)	64%(6), 86%(5)*
TARTU	Erdas (<i>ISODATA</i>)	Idrisi 1 (<i>rand. seed</i>)	Idrisi 2 (<i>rand. part.</i>)
ENVI	100%	76%(5)*	73% (6), 76%(5)*

*Ühtivus 5 /6 klassi korral

Maastikuindeksite arvutamiseks valiti ENVI tulemustega võrdlemiseks Kose jaoks 6 klassi korral ENVI tulemustega kõige sarnasem Idrisi klassifikatsioon – juhuslike keskpunktide (*random seed*) reeglga saadud tulemus, üldise ühtivusega 76% (joonis 5), ning Tartu jaoks Idrisi juhusliku jaotuse (*random partion*) reeglga saadud klassifikatsioon, mille klastrid 3 ja 4 liideti kokku nii ENVI kui Idrisi tulemuste puhul, saades üldiseks ühtivuseks 81% (joonis 6).



Joonis 5. k-means algoritmiga klassifitseeritud Kose testala tarkvarades **ENVI** (vasakul) ja **Idrisi** (*random seed*) (paremal).



Joonis 6. k-means algoritmiga klassifitseeritud Tartu testala tarkvarades **ENVI** (vasakul) ja **Idrisi** (*random partition*) (paremal).

Tabel 6. Algoritmi k-means klassifitseerimistulemuste ühtimine klasside lõikes Kose testalal Idrisi (*random seed*) ja ENVI tarkvarades (protsentides).

Idrisi ENVI	soine ala	vana mets	noor mets	põld1/ asustus	põld2/ infra.	rohumaa
soine ala	88.08	0	0	0	0.18	0
vana mets	11.92	85.29	0.89	0	0.05	0
noor mets	0	14.71	57.22	10.60	0	0
põld1/asustus	0	0	41.88	44.10	4.57	0
põld2/infra.	0	0	0	0.54	95.20	0.31
rohumaa	0	0	0	44.76	0	99.69

Tabel 7. Algoritmi k-means klassifitseerimistulemuste ühtimine klasside lõikes Tartu testalal Idrisi (*random partition*) ja ENVI tarkvarades (protsentides).

Idrisi ENVI	varjud	põld1/ mets1/ asustus	lage ala/ põld2/ mets2	põld3	roht- taimed
varjud	96.60	0	0	0	0
põld1/mets1/asustus	3.40	86.01	0	0	0
lage ala/põld2/mets2	0	13.99	75.46	0	0
põld3	0	0	24.54	61.08	0
rohttaimed	0	0	0	38.92	100.00

3.2. Etalonidega klassifitseerimine

Lühima tee ja suurima tõepära klassifikaatorite klassifitseerimise tulemusi võrreldi tarkvarades Erdas IMAGINE, ENVI, Idrisi Selva. Kasutades lühima tee algoritmi, ühtisid mõlemal testalal kõigi tarkvarade klassifitseerimistulemused üle 99% (tabel 8-9). Suurima tõepära algoritmi üldised klassifitseerimistulemused kõikusid veidi rohkem, kuni 5%, klasside lõikes erinesid tulemused kuni 12% (Tabel 10-11).

Tabel 8. Lühima tee klassifikaatori klassifitseerimistulemuste ühtivus Kose alal tarkvarades Erdas Imagine, ENVI, Idrisi Selva.

Minimum Distance	ENVI	Idrisi Selva
Erdas	99.43%	99.34%
ENVI		99.58%

Tabel 9. Lühima tee klassifikaatori klassifitseerimistulemuste ühtivus Tartu alal tarkvarades Erdas Imagine, ENVI, Idrisi Selva.

Minimum Distance	ENVI	Idrisi Selva
Erdas	99.60%	99.65%
ENVI		99.95%

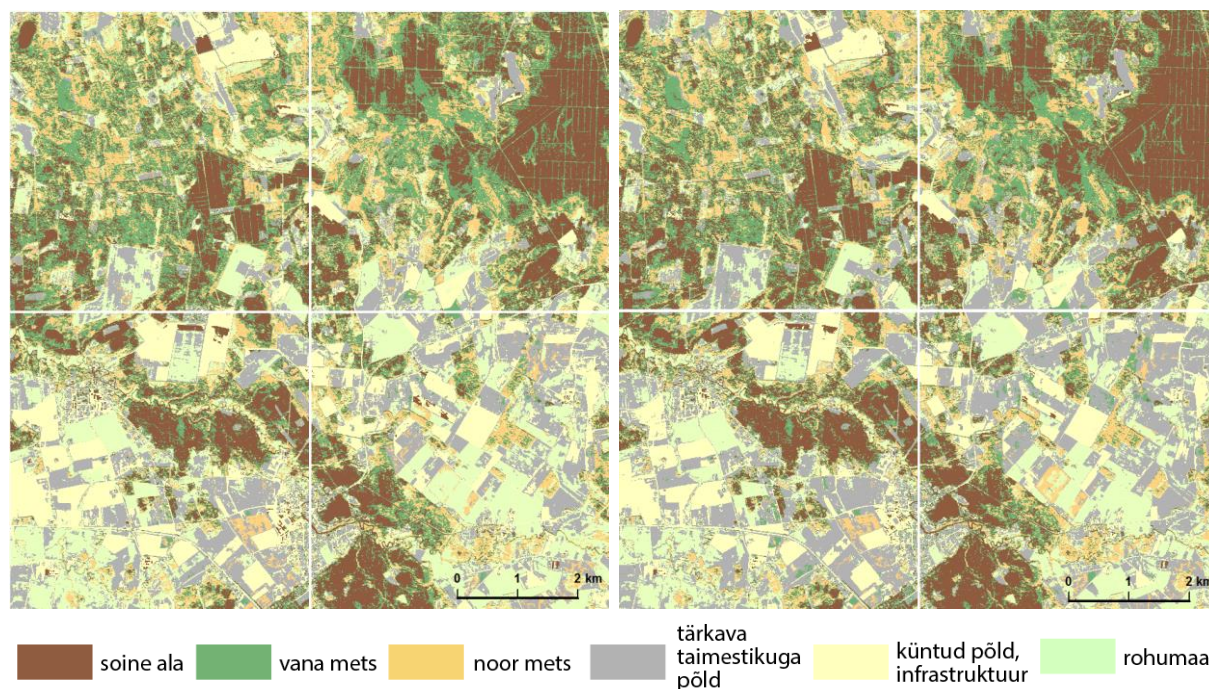
Tabel 10. Suurima tõepära klassifikaatori klassifitseerimistulemuste ühtivus Kose alal tarkvarades Erdas Imagine, ENVI, Idrisi Selva.

Maximum Likelihood	ENVI	Idrisi Selva
Erdas	94.99%	96.01%
ENVI		98.46%

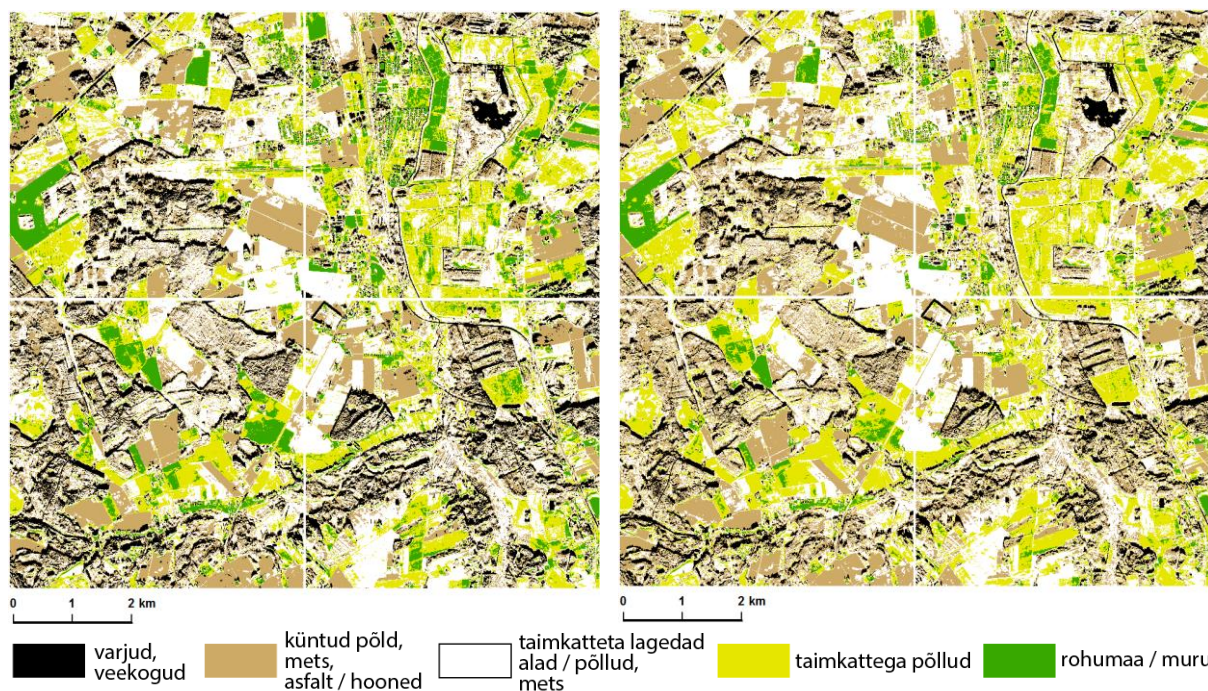
Tabel 11. Suurima tõepära klassifikaatori klassifitseerimistulemuste ühtivus Tartu alal tarkvarades Erdas Imagine, ENVI, Idrisi Selva.

Maximum Likelihood	ENVI	Idrisi Selva
Erdas	98.54%	98.37%
ENVI		99.82%

Näidistega klassifitseerimise tulemusi võrreldi ka ISODATA algoritmiga tehtud näidisteta klassifitseerimise tulemustega, kuna viimase põhjal valiti esimesele õpetuspiirkonnad, mistõttu need kaks klassifitseerimist peaksid andma sarnase tulemuse. Sobivaimaks näidistega klassifitseerimise klassifikaatoriks, võrdlemaks tulemusi ISODATA omadega, on lühima tee algoritm, kuna mõlemad põhinevad eukleidilise kauguse arvutamisel. Kose alal oli üldine ISODATA ja näidistega klassifitseerimise ühtivus **89,22%** (joonis 7) ning Tartu alal **88,8%** (joonis 8). Klasside lõikes olid erinevused suuremad. Kose alal tekkis erinevusi tärkava taimeistikuga põldude / asulate ning rohumaa klasside vahel, kus näidistega klassifitseerides oli põldude / asulate klassi ühilduvus näidisteta klassifitseerimisega 75% (tabel 12). Tartu alal tekkisid samad erinevused: näidisteta klassifitseerimisel saadi rohkem rohttaimede piksleid, võrreldes näidistega klassifitseerimisega (tabel 13).



Joonis 7. Tarkvaraga Erdas Imagine näidisteta (ISODATA) (vasakul) ja näidistega (*Minimum Distance*) (paremal) klassifitseeritud Kose testala.



Joonis 8. Tarkvaraga Erdas Imagine näidisteta (ISODATA) (vasakul) ja näidistega (*Minimum Distance*) (paremal) klassifitseeritud Tartu testala.

Tabel 12. Näidisteta (ISODATA) ja näidistega (*minimum distance*) klassifitseerimise tulemuste ühtivus klasside lõikes Kose testalal tarkvaraga Erdas IMAGINE (protsentides).

Näidisteta \ Näidistega	soine ala	mets1	mets2	põld1/ asustus	põld2/ infra.	rohumaa
soine ala	86.36	0.02	0	0	1.33	0
mets1	13.64	92.33	0.32	0	0.26	0
mets2	0	5.21	94.59	2.66	0	0
põld1/asustus	0	2.44	0	75.43	0.06	0
põld2/infra.	0	0	0	7.95	98.35	1.53
rohumaa	0	0	5.09	13.96	0	98.46

Tabel 13. Näidisteta (ISODATA) ja näidistega (*Minimum Distance*) klassifitseerimise tulemuste ühtivus klasside lõikes Tartu testalal tarkvaraga Erdas IMAGINE (protsentides).

Näidisteta \ Näidistega	varjud	põld1/ mets1/ asustus	lage ala/ põld2/ mets2	põld3	roht- taimed
varjud	100.00	14.49	0	0	0
põld1/mets1/asustus	0	83.62	2.05	0	0
lage ala/põld2/mets2	0	1.89	95.12	9.57	0
põld3	0	0	2.81	76.43	0
rohttaimed	0	0	0.03	14.00	100.00

3.3. Objektorienteeritud klassifitseerimine

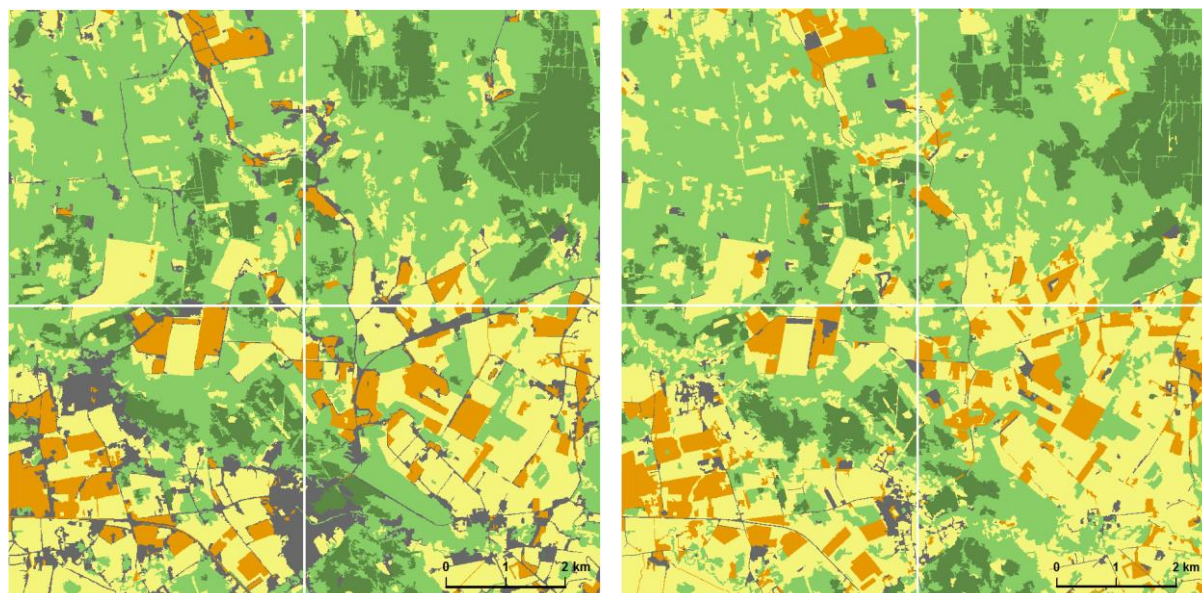
Üldine klassifitseerimise ühtivus tervel Kose alal oli 79%, kusjuures kõige enam ühtisid metsa iseloomustavad klassid ja kõige halvemini asustust kirjeldav maakasutuse klass (tabel 14; joonis 9), ning Tartu alal 75%, halvasti ühtisid siin asustuse ja varjude klassid, aga ka mets (tabel 15; joonis 10). Idrisi-ga klassifitseerides eristusid Tartu ala varjud halvasti, segunedes peamiselt asustusega. Varjude klass oli Idrisi tulemuse järgi palju väiksem kui eCognition-i põhjal.

Tabel 14. Tarkvarade eCognition ja Idrisi objektorienteeritud klassifitseerimise tulemuste ühtivus klasside lõikes Kose testalal (protsentides).

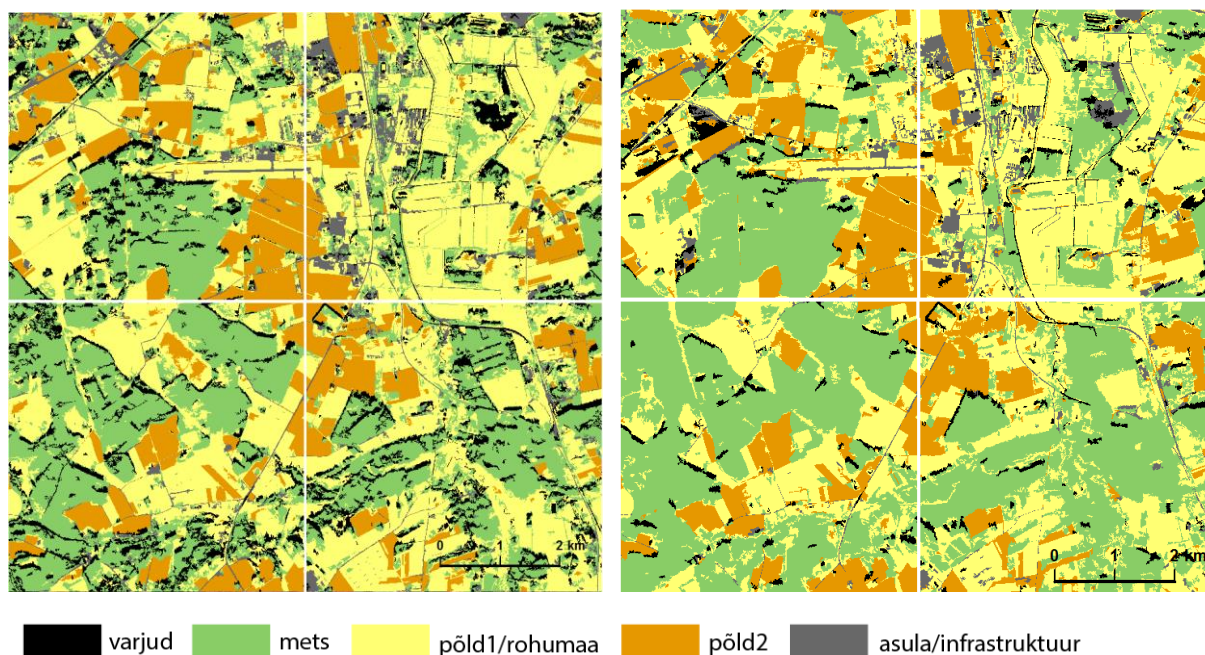
Idrisi eCog.	mets1	mets2	põld1	põld2	asustus
mets1	77.53	6.86	0.08	0.01	0.45
mets2	21.26	85.86	10.03	0.23	1.33
põld1	0.31	3.94	71.22	12.72	9.37
põld2	0.01	0.03	1.84	71.68	26.96
asustus	0.89	3.30	16.83	15.36	59.53

Tabel 15. Tarkvarade eCognition ja Idrisi objektorienteeritud klassifitseerimise tulemuste ühtivus klasside lõikes Tartu testalal (protsentides).

Idrisi eCog.	varjud	mets	põld1	põld2	asustus
varjud	59.82	16.46	1.07	0.41	20.24
mets	21.68	65.20	3.11	8.06	8.96
põld1	1.41	1.15	81.44	1.52	4.45
põld2	14.52	16.33	9.46	85.49	13.69
asustus	2.57	0.86	4.92	4.51	52.48



Joonis 9. Objektidepõhiselt klassifitseeritud Kose testala, kasutades eCognition (vasakul) ja Idrisi Selva (paremal) tarkvarasid.



Joonis 10. Objektidepõhiselt klassifitseeritud Tartu testala, kasutades eCognition (vasakul) ja Idrisi Selva (paremal) tarkvarasid.

3.4. Klassifitseerimistulemuste põhjal arvutatud maastikuindeksite varieerumine

Eri programmides või eri viisil klassifitseeritud satelliidipiltidelt arvutatud maastiku taseme indeksite varieeruvusi võrreldi kahel 100 km² testalal (RapidEye Kose ümbrust kujutav satelliidipilt ja GeoEye Tartu linnast lõunasse jäävat ala kujutav satelliidipilt) ning nende neljaks väiksemaks alaks jaotamisel saadud kaheksal testalal. Töötlusviisid, mille mõju maastikumeetrikale hinnati, olid näidisteta klassifitseerimine eri tarkvarades (algoritmid ISODATA ja k-means), näidistega ja näidisteta klassifitseerimise võrdlus ning objektorienteeritud klassifitseerimine tarkvarades eCognition ning Idrisi Selva. Maastikuindeksite arvutamisel näidisteta klassifitseeritud piltidelt kasutati vaid neid töötlusi, kus algoritm oli klasterdanud pikslid korduval katsetamisel ühtemoodi.

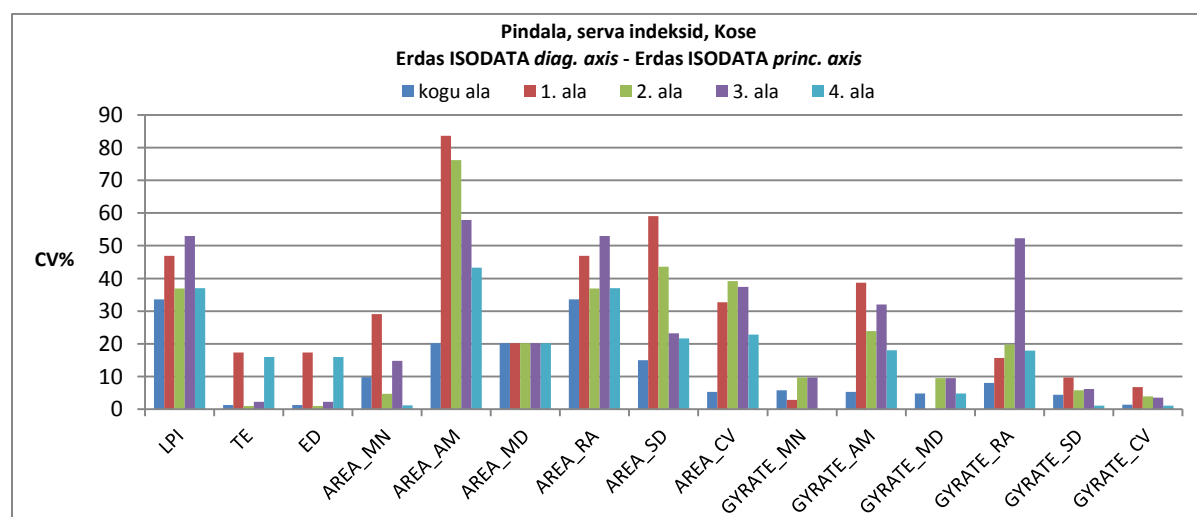
Pindala ja serva indeksid

Pindala ja serva indeksitest varieerusid kõige vähem pöörlemise raadiust ehk eraldiste ulatust / kompaktsust mõõtvad indeksid GYRATE_MN (keskmine), GYRATE_MD (mediaan), GYRATE_CV (variatsioonikoeffitsent) (tabel 16; joonis 11-12). Kõige enam varieerusid suurima eraldise indeks LPI, kõik eraldiste pindalalist jaotumist kirjeldavad indeksid (pindalaga kaalutud keskmine pindala AREA_AM, pindalade ulatus AREA_RA, mediaan pindala AREA_MD, pindalade standardhälve AREA_SD, pindalade variatsioonikoeffitsent AREA_CV) ning pöörlemise raadiust ehk eraldiste ulatust mõõtvatest indeksitest GYRATE_AM (pindalaga kaalutud keskmine – suurematele pindaladel antakse suurem kaal) ja GYRATE_RA. Kuna GYRATE_AM ja GYRATE_RA sõltuvad rohkem pindalade (AREA) varieeruvusest kui GYRATE keskmisi väärtusi kirjeldavad indeksid (MN, SD, CV), siis suurendas see ka GYRATE_AM ja GYRATE_RA variatsioonikoeffitsente. LPI, mis näitab suurima eraldise protsentuaalset osakaalu maastikus, variatsioonikoeffitsendid varieerusid üsna palju alade lõikes. Keskmise stabiilsusega indeksid olid kogu serva pikkus TE, servatihedus ED, pöörlemise raadiuse ehk ulatuse standardhälve GYRATE_SD ning

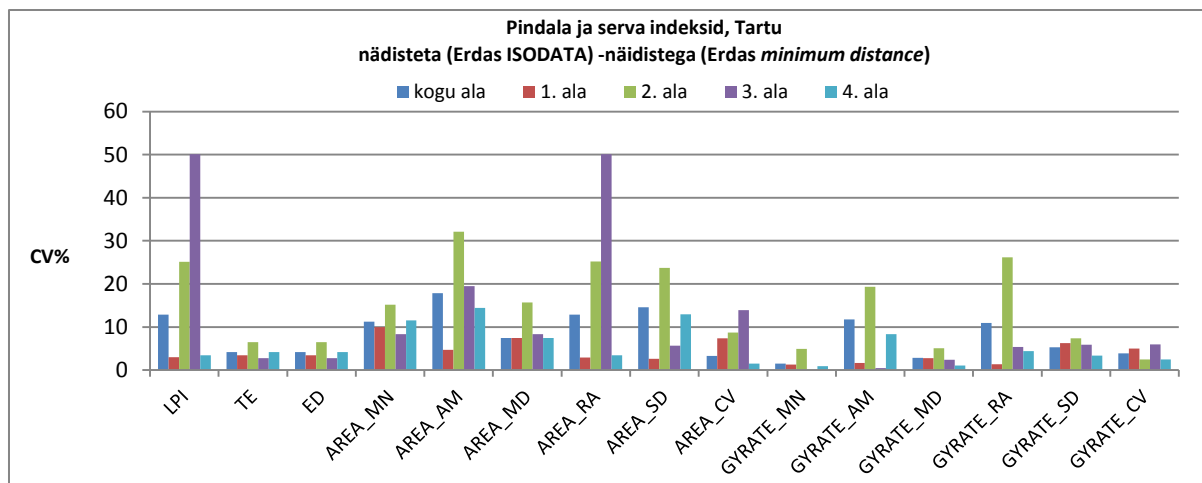
pindalade keskmine AREA_MN, mille keskmised variatsioonikoeffitsendid ei olnud küll eriti suured, kuid mis alade või töötlusviiside lõikes erinesid üsna palju. Kuna servatihedus sõltub servapikkusest, siis TE ja ED korreleeruvad täielikult ning nende variatsioonikoeffitsendid tulid kõikide võrdluspaaride jaoks samad. Rohkem varieerusid TE ja ED väärtused objektorienteeritud klassifitseerimiste võrdluses Tartu alal, kus Idrisi-ga klassifitseerides eristus vähem eraldisi (joonis 13). Eelkõige ei eristunud eCognition-iga võrreldes nii hästi varjude klass. Kuna TE ja ED sõltuvad eraldiste arvust, siis korreleeruvad need indeksid väga hästi ka koonduvuse tüübi alla kuuluvate indeksite NP (eraldiste arv) ning PD-ga (eraldiste tihedus).

Tabel 16. Pindala ja serva indeksite keskmised variatsioonikoeffitsendid alade ja töötlusviiside kaupa järjestatud kasvavalt kõikide alade keskmise järgi.

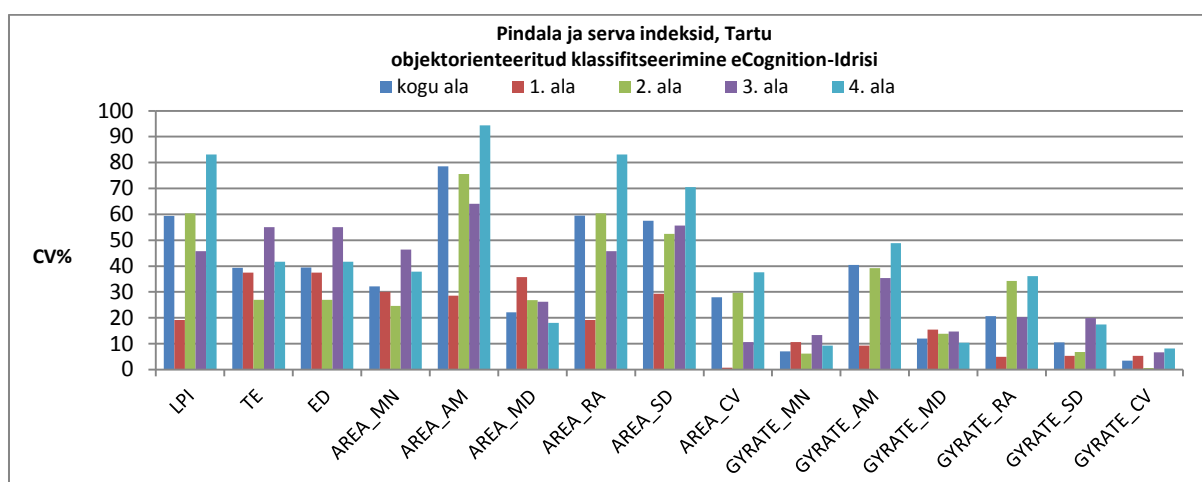
	Kose					Tartu				
	ISODATA Erdas	ISODATA Erdas vs Idrisi	k- means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	k- means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	kõik alad keskmine	Kõik alad mediaan
GYRATE_CV	3.36	2.18	4.86	2.34	2.24	5.41	3.94	4.85	3.65	3.75
GYRATE_MN	5.64	3.67	3.98	0.97	11.70	1.00	1.73	9.32	4.75	3.68
GYRATE_MD	5.73	2.89	8.04	0.00	10.15	2.66	2.82	13.30	5.70	4.53
GYRATE_SD	5.47	2.56	1.64	2.94	12.34	6.41	5.64	12.00	6.12	5.31
TE	7.59	2.67	4.53	0.82	10.52	8.41	4.19	40.09	9.85	5.11
ED	7.59	2.67	4.58	0.82	10.52	8.43	4.22	40.09	9.87	5.11
AREA_MN	11.94	7.47	1.93	1.59	12.04	14.60	11.24	34.20	11.88	9.48
GYRATE_RA	22.77	5.10	6.44	21.73	9.84	17.04	9.65	23.26	14.48	10.44
AREA_MD	20.20	16.16	20.20	0.00	27.72	6.30	9.27	25.77	15.70	19.73
AREA_CV	27.51	7.85	8.29	16.50	20.07	24.55	6.96	21.35	16.64	10.97
GYRATE_AM	23.62	8.15	8.41	15.81	10.48	30.42	8.30	34.66	17.48	16.00
AREA_SD	32.52	12.22	8.10	17.67	14.45	38.36	11.90	53.06	23.53	19.43
AREA_RA	41.50	17.72	20.69	36.75	23.95	40.17	18.89	53.58	31.65	28.86
LPI	41.50	17.72	20.75	36.75	23.95	40.15	18.90	53.57	31.66	28.91
AREA_AM	56.23	18.94	16.20	32.84	28.92	58.76	17.70	68.23	37.23	30.32



Joonis 11. Pindala ja serva indeksite variatsioonikoeffitsendid (CV%) Erdas ISODATA algoritmiga klassifitseerimisel, kasutades diagonaalse telje (*diagonal axis*) ja peakomponentide telje (*principal axis*) klastrite algatamise reegleid Kose alal alade kaupa.

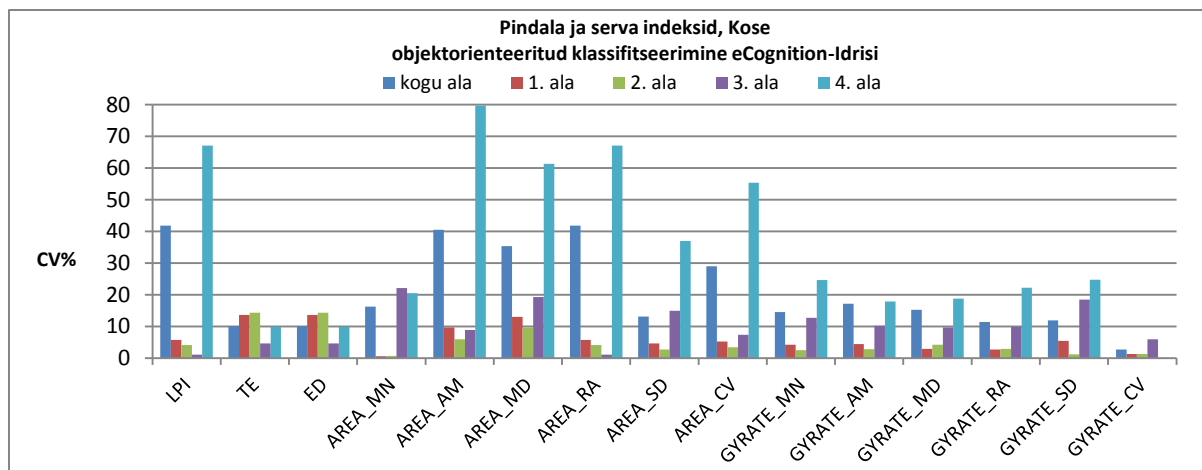


Joonis 12. Pindala ja serva indeksite variatsioonikoeffitsendid (CV%) nädisteta (Erdas ISODATA) ja nädistega (Erdas *Minimum Distance*) klassifitseerimisel, Tartu alade kaupa.



Joonis 13. Pindala ja serva indeksite variatsioonikoeffitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Tartu alade kaupa.

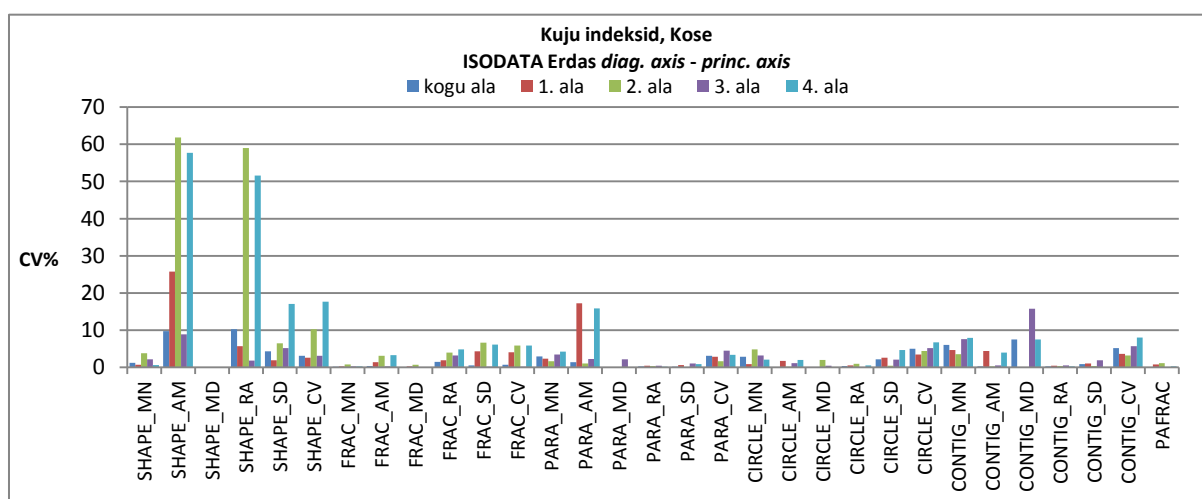
Alade lõikes eristus objektorienteeritud klassifitseerimisega saadud pindala ja serva indeksite võrdlemisel Kose 4. ala, mille variatsioonikoeffitsendid olid kõrgemad kui teistel (joonis 14). Põhjuseks on ilmselt teed ja asulad, mis Idrisi-ga klassifitseerides ei eristunud, kuid eristusid eCognition-iga, muutes tunduvalt maastiku struktuuri.



Joonis 14. Pindala ja serva indeksite variatsioonikoeffitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Tartu alal alade kaupa.

Kuju indeksid

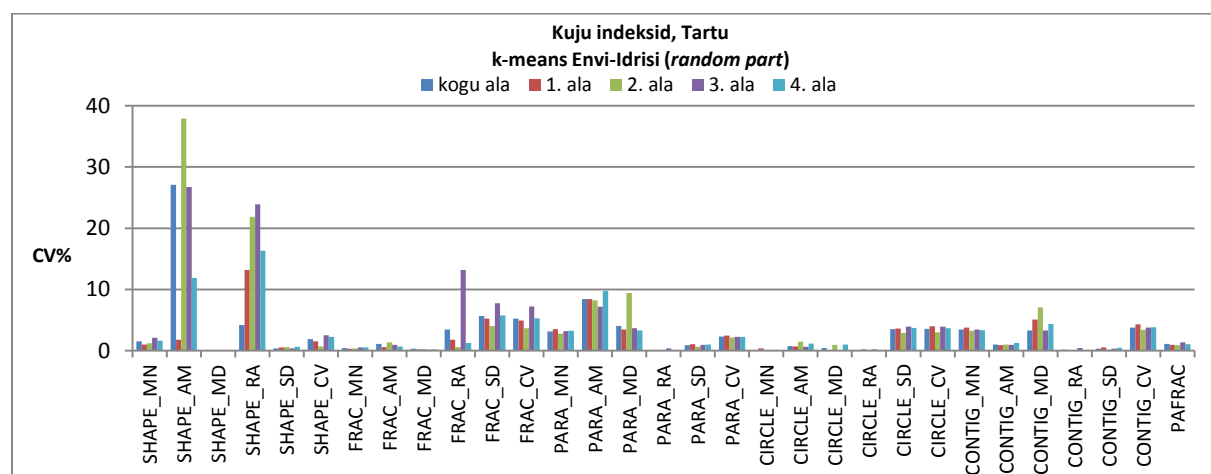
Kuju indeksite variatsioonikoeffitsendid ei olnud üldiselt eriti suured eri töötluste ja alade lõikes, jäädes enamasti alla 10% (tabel 17; joonis 15-16). Erandiks olid vaid pindalaga kaalutud kuju indeks SHAPE_AM ja kuju indeksi väärtuste ulatus SHAPE_RA, mis varieerusid kõige rohkem. SHAPE_AM sõltub eraldiste pindalade suurusest, mis varieerusid üsna palju, ning SHAPE_RA varieeruvust suurendab see, et indeksi SHAPE võimalik väärtuste vahemik on 1 kuni lõpmatus, samas kui teiste kuju indeksite oma jääb kindlasse vahemikku (0 kuni 1 või 1 kuni 2). Pooled kuju indeksitest olid stabiilsed (CV% alla 10) kõikide alade ja töötlusviiside lõikes: SHAPE_MD, PARA_RA, CONTIG_RA, CONTIG_AM, FRAC_MD, FRAC_MN, FRAC_AM, peaaegu kõik väikseima võimaliku ümbritseva ringi indeksid (CIRCLE_RA, CIRCLE_AM, CIRCLE_MD, CIRCLE_MN, CIRCLE_SD).



Joonis 15. Kuju indeksite variatsioonikoeffitsendid (CV%) Erdas ISODATA algoritmiga klassifitseerimisel, kasutades diagonaalse telje (*diagonal axis*) ja peakomponentide telje (*principal axis*) klastrite algatamise reegleid Kose alal alade kaupa.

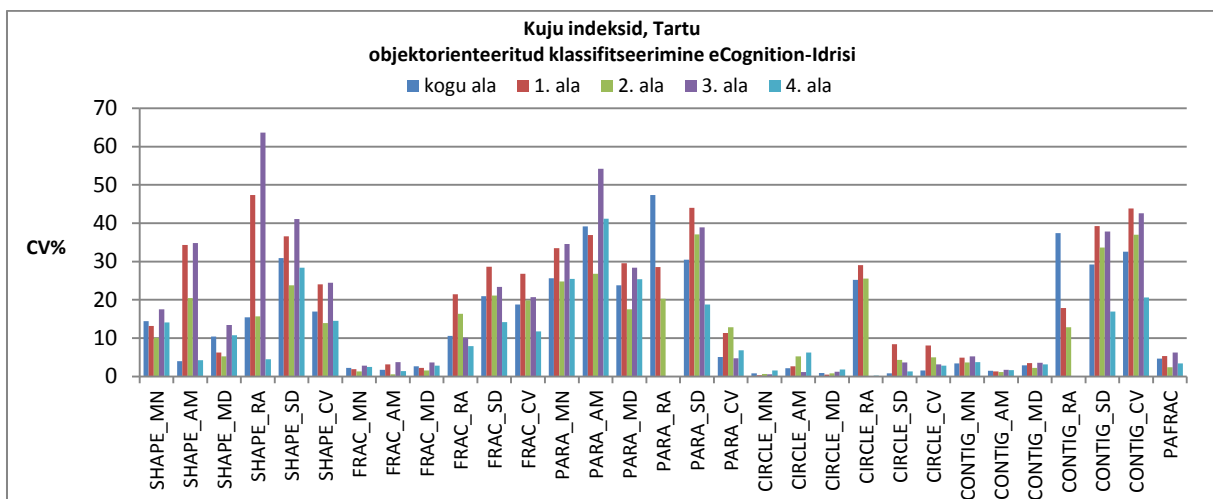
Tabel 17. Kuju indeksite keskmised variatsioonikoefitsendid töötlusviiside kaupa järjestatud kasvavalt kõikide alade keskmise järgi.

	Kose					Tartu				
	ISODATA Erdas	ISODATA Erdas vs Idrisi	k- means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	k- means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	kõik alad keskmine	kõik alad mediaan
FRAC_MD	0.18	0.13	0.18	0.14	0.40	0.23	0.25	2.57	0.51	0.24
FRAC_MN	0.34	0.16	0.51	0.11	0.40	0.42	0.54	2.13	0.58	0.34
CIRCLE_MD	0.47	0.00	1.09	0.03	1.25	0.46	0.51	1.04	0.61	0.40
CONTIG_AM	1.88	0.61	1.11	0.22	0.47	1.02	0.52	1.46	0.91	0.58
FRAC_AM	1.68	0.21	0.58	0.29	1.09	0.94	0.56	2.11	0.93	0.62
CIRCLE_AM	1.03	0.66	0.57	0.40	3.07	0.93	0.79	3.48	1.36	0.74
CIRCLE_MN	2.76	2.22	3.84	0.30	0.60	0.14	0.44	0.78	1.39	0.62
SHAPE_MD	0.00	0.00	0.00	0.00	3.46	0.00	0.16	9.21	1.60	0.00
PAFRAC	0.45	0.15	1.49	0.42	4.15	1.06	1.04	4.41	1.64	1.09
CONTIG_RA	0.33	0.08	0.27	0.32	3.97	0.15	0.08	13.63	2.35	0.17
CIRCLE_SD	2.36	1.73	1.77	0.38	2.09	3.52	4.62	3.70	2.52	2.08
CIRCLE_RA	0.48	0.24	0.27	0.51	5.37	0.11	0.08	16.02	2.88	0.27
PARA_RA	0.31	0.07	0.25	0.28	7.05	0.12	0.08	19.24	3.43	0.13
FRAC_RA	3.05	1.31	0.91	1.05	4.49	4.06	0.32	13.28	3.56	1.56
CONTIG_MN	5.95	4.56	4.80	0.89	0.79	3.45	4.28	4.20	3.61	3.76
CIRCLE_CV	4.94	3.95	5.61	0.57	2.31	3.63	4.76	4.12	3.74	4.01
SHAPE_MN	1.67	0.97	2.74	0.41	7.59	1.50	1.68	13.84	3.80	1.74
PARA_CV	3.10	2.33	0.97	0.81	11.51	2.27	2.56	8.15	3.96	2.46
CONTIG_MD	6.12	6.12	6.49	2.24	1.02	4.61	5.50	3.06	4.40	3.29
CONTIG_SD	0.84	0.62	1.06	0.58	10.06	0.36	0.73	31.38	5.70	0.85
PARA_MD	0.44	0.44	0.44	0.57	9.88	4.78	5.56	24.92	5.88	2.73
PARA_MN	2.93	2.18	2.32	0.44	6.00	3.17	4.03	28.77	6.23	2.80
FRAC_CV	3.33	0.84	3.10	1.18	9.74	5.27	6.88	19.56	6.24	4.47
PARA_SD	0.54	0.28	1.35	0.49	11.71	0.90	1.47	33.84	6.32	1.02
FRAC_SD	3.55	0.78	3.61	1.28	10.07	5.69	7.41	21.64	6.75	4.87
SHAPE_CV	7.31	0.91	2.65	1.31	19.95	1.76	3.64	18.77	7.04	2.89
CONTIG_CV	5.13	3.94	5.86	0.42	9.93	3.80	5.02	35.33	8.68	4.92
SHAPE_SD	6.98	1.85	5.39	1.07	27.32	0.49	1.96	32.15	9.65	2.81
PARA_AM	7.55	2.65	4.56	0.82	10.24	8.40	4.20	39.64	9.76	5.05
SHAPE_AM	32.78	8.27	7.40	13.54	13.89	21.09	9.32	19.56	15.73	11.72
SHAPE_RA	25.65	12.19	13.02	13.38	31.94	15.90	7.08	29.31	18.56	13.30



Joonis 16. Kuju indeksite variatsioonikoefitsendid (CV%) k-means algoritmiga klassifitseerimisel tarkvarades ENVI ja Idrisi (random partition).

Suuremad indeksite variatsioonikoeffitsendid saadi objektorienteeritud klassifitseerimistulemuste põhjal arvutatud kuju indeksite väärtuste võrdlemisel, eriti Tartu aladel (joonis 17). Tartu alal on ilmselt kuju varieeruvuse põhjuseks jällegi varjud, mis Idrisi-ga ei avaldunud eriti, kuid mille kuju on üsna keerukas. Teisalt oli siin ka indekseid, mis varieerusid vähe, nagu übermõõdu-pindala fraktaalne dimensioon PAFRAC ja väikseima võimaliku ümbritseva ringi indeksid CIRCLE_. Lisaks varieerusid Tartu aladel vähe fraktaalset dimensiooni (FRAC_) ja pidevust / kuju kompaktsust (CONTIG_) kirjeldavate indeksite keskmised väärtused, kuna FRAC ja CONTIG väärtused jäävad kitsasse vahemikku. PAFRAC ja FRAC väärtused ei sõltu niivõrd pindalast ega übermõõdust, nagu näiteks indeksite rühma PARA_ või SHAPE_ väärtused, kuna kasutavad übermõõdu ja pindala logaritme.



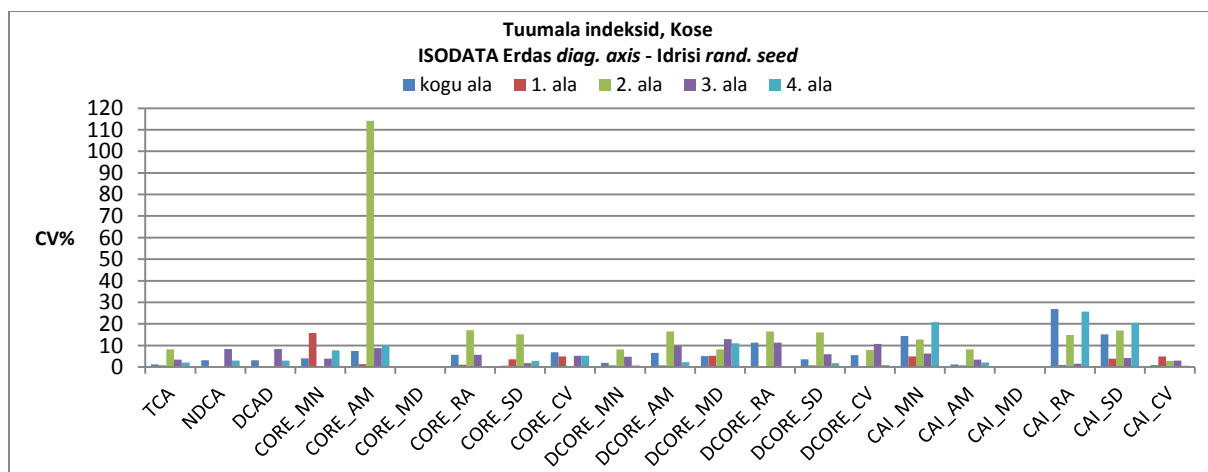
Joonis 17. Kuju indeksite variatsioonikoeffitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Tartu alal alade kaupa.

Tuumala indeksid

Tuumala indeksite variatsioonikoeffitsendid erinesid eri töötlusviiside ja alade lõikes üsna palju ning kõikide töötlusviiside ja alade lõikes stabiilseid indekseid, mille CV% jäänuks alla 10, peaaegu ei olnudki (tabel 18). Väga suurt stabiilsust näitasid siiski kaks indeksit – tuumalade pindalate mediaan CORE_MD ning tuumalade osakaalude mediaan eraldistes CAI_MD, mille varieeruvuskoeffitsendid olid kõikide töötluste ning alade lõikes 0-id. Põhjuseks on ilmselt paljude väikeste või piklike eraldiste suur osakaal maastikus, mistõttu autori määratud servasügavuste juures (Kose 100 m ja Tartu 50 m) ei olnudki paljudel eraldistel tuumalaid. Keskmise stabiilsusega olid variatsioonikoeffitsenti kirjeldavad indeksid (CAI_CV, DCORE_CV, CORE_CV) ning liigendatud tuumalade arv NDCA ja liigendatud tuumalade tihedus DCAD. DCAD sõltub NDCA-st, mistõttu need indeksid korreleeruvad täielikult, mida näitasid ka variatsioonikoeffitsendid. Oli ka võrdluspaare, kus varieeruvus oli enamike indeksite puhul väga väike, nagu Kose aladel ISODATA algoritmi võrdlemisel tarkvarades Erdas ja Idrisi (joonis 18). Siiski varieerus märgatavalt 2. ala pindalaga kaalutud tuumala indeks CORE_AM, ka selle sama ala ja töötluste pindalaga kaalutud pindala AREA_AM variatsioonikoeffitsent oli kõrge.

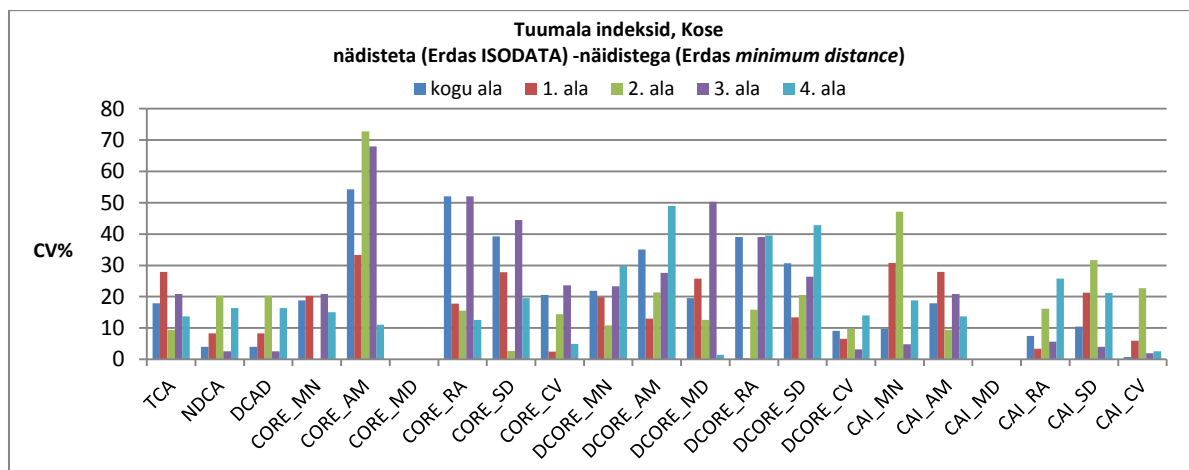
Tabel 18. Tuumala indeksite keskmised variatsioonikoefitsendid töötlusviiside kaupa järjestatud kasvavalt kõikide alade keskmise järgi.

	Kose					Tartu				
	ISODATA Erdas	ISODATA Erdas vs Idrisi	k- means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	k- means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	kõik alad keskmine	kõik alad mediaan
CORE_MD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CAI_MD	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CAI_CV	12.40	2.50	8.90	6.80	16.30	5.30	12.70	11.00	9.47	7.09
DCORE_CV	16.70	5.00	7.30	8.50	13.40	12.50	4.10	29.40	12.11	8.85
DCAD	33.90	2.90	8.60	10.30	8.20	16.80	5.30	16.10	12.76	8.70
NDCA	33.90	2.90	8.50	10.30	8.20	16.90	5.40	16.10	12.76	8.72
CORE_CV	23.90	4.50	6.10	13.20	20.30	8.40	9.50	17.50	12.93	8.19
CAI_RA	49.70	14.00	9.30	11.70	14.00	10.10	18.90	5.90	16.68	10.88
CAI_SD	54.50	12.10	14.10	17.70	19.10	4.20	10.30	17.10	18.63	12.88
CAI_AM	84.30	3.20	13.70	18.00	5.50	17.20	10.20	31.00	22.87	12.97
TCA	84.30	3.20	13.70	18.00	5.50	17.30	10.20	31.00	22.88	12.99
DCORE_MN	62.20	3.30	16.80	21.10	13.20	12.40	13.80	45.40	23.52	16.56
CAI_MN	61.80	11.80	18.30	22.20	33.00	8.70	13.30	27.70	24.60	17.06
DCORE_SD	70.70	5.60	12.30	26.80	15.30	17.90	12.10	68.20	28.62	15.93
CORE_MN	75.20	6.30	9.10	15.00	15.00	30.20	19.90	61.30	28.99	19.60
DCORE_MD	77.20	8.40	34.50	21.90	11.80	19.50	42.50	30.70	30.83	19.56
DCORE_RA	74.90	7.90	6.50	26.80	27.30	26.60	18.40	68.10	32.05	18.35
DCORE_AM	72.40	7.20	10.90	29.20	25.20	24.50	12.20	82.90	33.06	17.49
CORE_SD	86.40	4.80	9.60	26.70	17.50	39.00	18.60	73.40	34.51	25.46
CORE_RA	93.70	6.00	4.50	30.00	29.70	41.90	17.30	77.00	37.51	29.38
CORE_AM	102.20	28.40	32.40	47.90	34.30	75.50	28.50	86.90	54.50	50.61

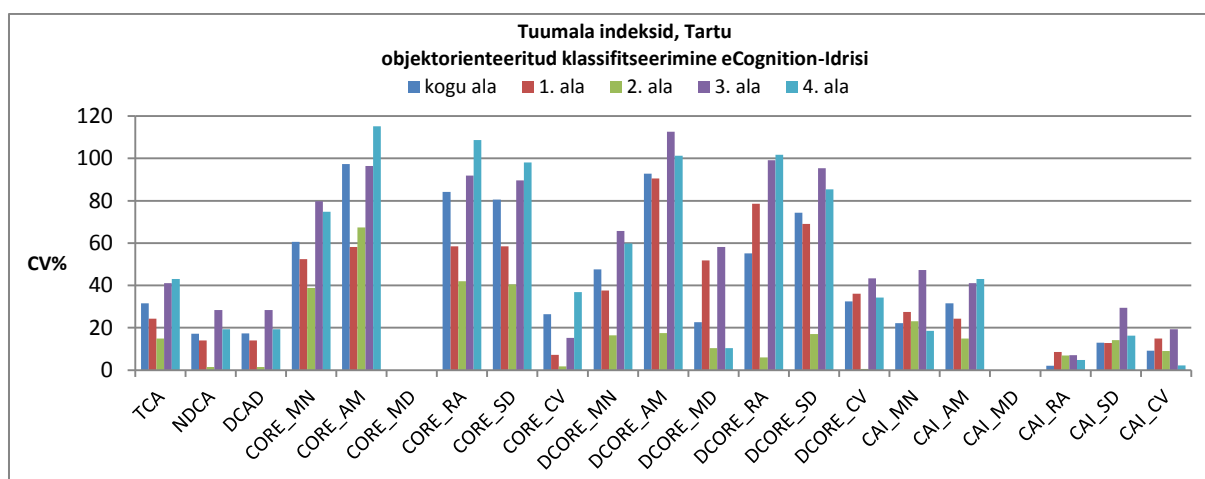


Joonis 18. Tuumala indeksite variatsioonikoefitsendid (CV%) ISODATA algoritmiga klassifitseerimisel tarkvarades Erdas (*diagonal axis*) ja Idrisi (*random seed*) Kose alal alade kaupa.

Enamike alade ning töötlusviiside võrdluspaaride puhul tulid aga üsna erinevad variatsioonikoefitsendid, nagu Kose nädisteta ja nädistega klassifitseerimiste võrdlemisel (joonis 19) või Tartu objektorienteeritud klassifitseerimiste võrdlemisel (joonis 20), mis kinnitab tuumala indeksite ebastabiilsust



Joonis 19. Tuumala indeksite variatsioonikoefitsendid (CV%) näidisteta (Erdas ISODATA) ja näidistega (Erdas *Minimum Distance*) klassifitseerimisel, Kose alal alade kaupa.



Joonis 20. Tuumala indeksite variatsioonikoefitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Tartu alal alade kaupa.

Koonduvuse indeksid

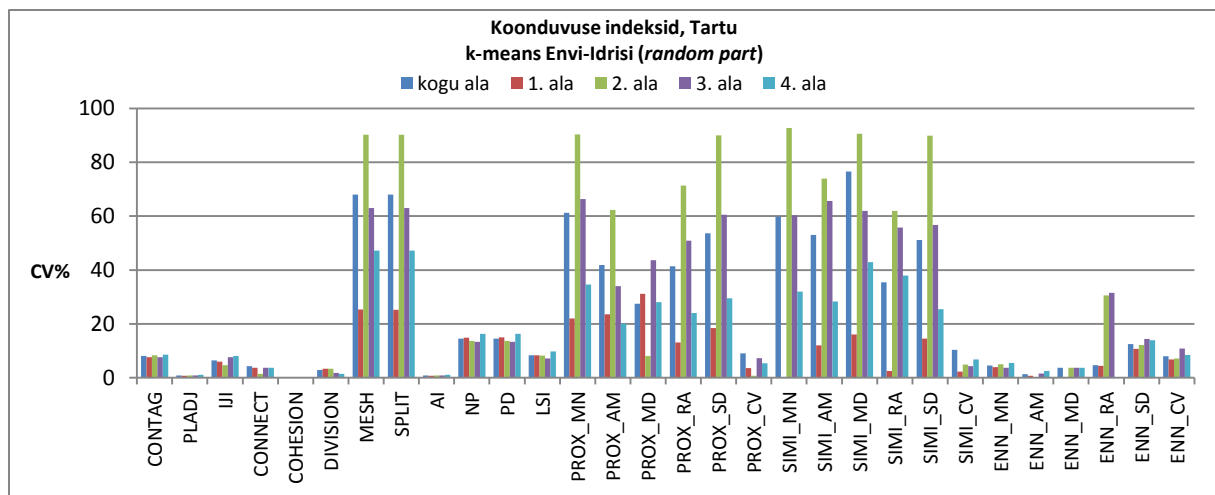
Kolmandik koonduvuse indeksitest olid stabiilsed kõikidel aladel ja tarkvaratöötluste suhtes: eraldiste sidususe indeks COHESION, agregatsiooni / ühendatuse indeks AI, sama tüüpi naabrussuhete protsenti mõõtev indeks PLADJ, maastiku eraldisteks jagunemist kirjeldav indeks DIVISION, sidususe indeks CONNECT, koonduvuse indeks CONTAG, lähima naabri indeksid ENN_AM ja ENN_CV (tabel 19; joonis 21-22). Enamikul neist indeksitest on kindel väärtuste vahemik (arvutatakse protsentides, DIVISION-i väärtuste vahemik on 0 kuni 1), välja arvatud ENN_, millel ei ole piirväärtust, ning COHESION, mille ulatus ja käitumine maastiku tasemel ei ole teada. Väga suuri varieeruvusi ei esinenud ka vahelduvuses / kõrvuti paiknemises IJI, maastiku kuju indeksis LSI, lähima naabri indeksites (ENN_MD, ENN_MN, ENN_SD, ENN_RA) eraldiste arvus NP, eraldiste tiheduses PD. Ebastabiilsed olid indeksid MESH (koondatus) ja SPLIT (tükeldatus) ning suuremaid varieeruvusi esines läheduse (PROX_) ja sarnasuse (SIMI_) indeksite rühmas, kus variatsioonikoefitsendid erinesid üsna palju ka Tartu või Kose väiksemate alade sees, küündides mõnel alal üle 100%, teisel alal jälle jäädes vaid 10%-ni. Indeksitel PROX_

ja SIMI_ ei ole ühikuid ega piirväärtust ning SIMI_ on indeksi PROX_ modifikatsioon, võttes lisaks eraldiste suurusele ja lähedusele arvesse ka sarnasuse maatriksiga määratud maastikuklasside / eraldiste üksteisega sarnanemist. Kuna mõlemad indeksid põhinevad eraldiste pindalal, siis on üheks ebastabiilsuse põhjuseks ilmselt pindalade suur varieeruvus (vt pindala ja serva indeksid), kuid SIMI_ indeksite varieeruvust suurendab veelgi sõltuvus sarnasuse maatriksist.

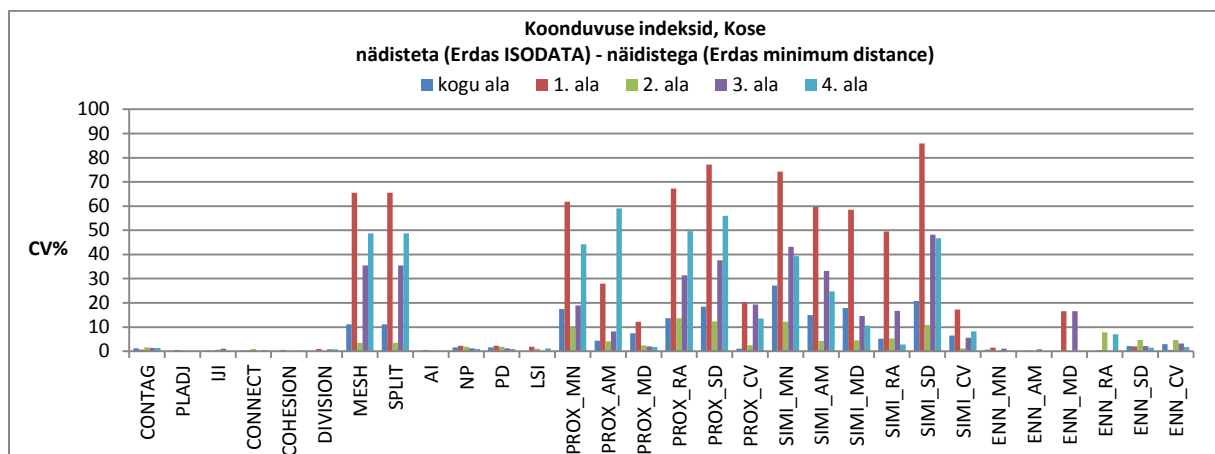
Indeksid MESH ja SPLIT varieerusid samade töötlusmeetodite korral ühtemoodi, mis on ka loogiline, kuna need indeksid on täpselt teineteise vastandid ning korreleeruvad täielikult. Lisaks korreleeruvad need indeksid täielikult ka indeksiga pindalaga kaalutud keskmine pindala AREA_AM – varieeruvuskoeffitsendid olid ühesuguse töötlusmeetodi ja sama ala korral samaväärsed. Põhjuseks on see, et SPLIT ja MESH põhinevad eraldiste pindalal.

Tabel 19. Koonduvuse indeksite keskmised variatsioonikoeffitsendid töötlusviiside kaupa järjestatud kasvavalt kõikide alade keskmise järgi.

	Kose					Tartu				
	ISODATA Erdas	ISODATA Erdas vs Idrisi	k-means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	k-means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	kõik alad keskmine	kõik alad mediaan
COHESION	0.48	0.10	0.10	0.23	0.06	0.14	0.07	0.10	0.16	0.13
AI	1.65	0.57	1.00	0.19	0.35	0.90	0.47	1.24	0.80	0.53
PLADJ	1.65	0.57	1.00	0.19	0.36	0.91	0.47	1.24	0.80	0.53
DIVISION	1.73	0.70	0.60	0.54	3.52	2.54	0.41	5.38	1.93	0.86
CONNECT	2.58	0.57	1.44	0.44	4.57	3.39	1.43	14.18	3.57	1.62
IJI	11.87	0.88	4.92	0.43	3.38	6.56	4.46	3.14	4.45	3.58
ENN_AM	2.15	0.82	0.67	0.28	9.47	1.26	0.88	20.96	4.56	1.12
CONTAG	7.45	2.01	2.29	1.23	4.74	8.04	3.67	7.61	4.63	3.45
ENN_CV	7.37	1.31	1.87	2.59	6.83	8.24	2.04	7.74	4.75	3.22
ENN_MN	2.04	1.38	0.85	0.72	10.18	4.55	5.21	38.15	7.88	3.85
LSI	7.51	2.64	4.51	0.81	9.97	8.35	4.17	39.20	9.65	5.03
ENN_MD	0.00	0.00	3.31	6.62	15.72	2.98	8.53	41.69	9.86	3.72
SIMI_CV	20.40	9.78	11.02	7.74	4.60	5.71	10.11	9.56	9.87	7.02
ENN_SD	9.09	2.69	1.49	2.47	14.43	12.77	7.25	30.83	10.13	6.14
PROX_CV	18.15	5.79	6.18	11.30	12.67	5.22	14.23	12.14	10.71	8.97
NP	11.93	7.46	1.89	1.58	12.04	14.56	11.21	34.19	11.86	9.47
PD	11.93	7.46	1.92	1.58	12.04	14.59	11.23	34.20	11.87	9.47
ENN_RA	34.97	8.71	8.70	3.07	4.78	14.26	4.31	23.52	12.79	7.32
PROX_MD	40.28	8.11	6.59	5.17	45.73	27.68	21.72	51.64	25.87	21.51
SIMI_RA	38.54	15.26	9.13	15.89	29.71	38.70	20.64	41.00	26.11	18.75
PROX_RA	48.71	20.95	13.25	35.09	29.79	40.15	19.66	42.40	31.25	26.89
PROX_MN	61.25	19.15	17.00	30.45	9.25	54.91	26.10	46.71	33.10	27.42
SIMI_AM	54.83	17.74	13.06	27.38	37.78	46.60	19.02	49.56	33.25	23.45
PROX_AM	49.01	37.80	25.06	20.75	31.37	36.39	29.49	49.35	34.90	27.58
PROX_SD	62.35	18.44	17.75	40.31	18.55	50.41	24.02	47.91	34.97	29.17
SIMI_SD	55.44	19.27	21.54	42.50	30.04	47.54	22.70	57.27	37.04	34.54
MESH	56.23	18.94	16.20	32.84	28.92	58.76	17.70	68.23	37.23	30.32
SPLIT	56.23	18.94	16.26	32.84	28.92	58.74	17.71	68.23	37.23	30.31
SIMI_MN	63.03	22.03	17.23	39.19	30.12	48.97	13.95	64.35	37.36	32.43
SIMI_MD	56.03	24.87	17.17	21.21	34.28	57.64	38.41	75.95	40.70	40.74



Joonis 21. Koonduvuse indeksite variatsioonikoefitsendid (CV%) k-means algoritmiga klassifitseerimisel tarkvarades ENVI ja Idrisi (*random partition*) Tartu alal alade kaupa.



Joonis 22. Koonduvuse indeksite variatsioonikoefitsendid (CV%) näidistega ja näidisteta klassifitseerimisel Kose alal alade kaupa.

Kontrastsuse indeksid

Kontrastsuse indeksid ei varieerunud üldiselt eriti palju. Kuna need on funktsionaalsed indeksid, mille väärtused sõltuvad kasutaja poolt määratud serva kontrastsusest (määratakse võttes arvesse uuritavat protsessi), siis erinevad serva kontrastsuse maatriksid võivad anda erinevaid tulemusi.

Stabiilsed indeksid olid selles rühmas peaaegu kõik serva kontrastsuse jagunemist maastikus kirjeldavad indeksid ECON_, vaid pindalaga kaalutud keskmine indeks ECON_AM varieerus mõnel alal rohkem, nagu ka teised kaks kontrastsuse indeksit – kogu serva kontrastsuse indeks TECI ja kontrastsusega kaalutud servade tihedus CWED (tabel 20).

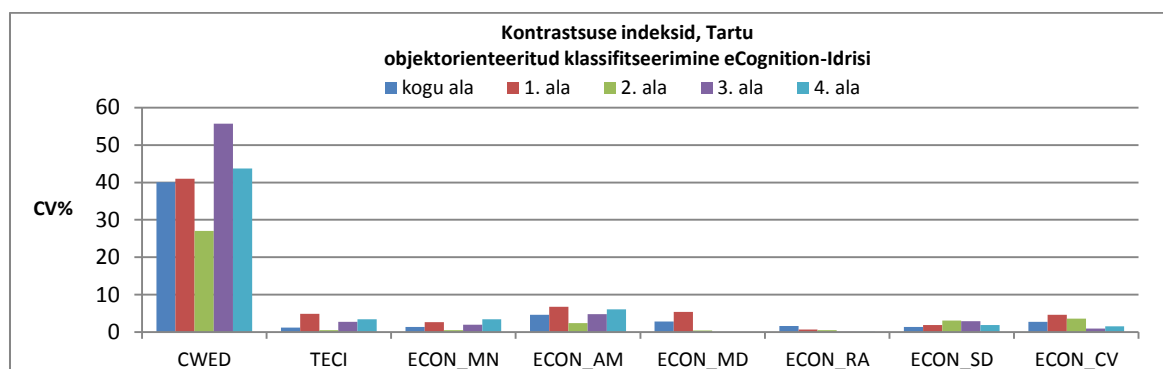
Tartu ala objektorienteeritud klassifitseerimisel saadud serva kontrastsuse indeks TECI ja serva kontrastsuse jagunemist maastikus kirjeldavad indeksid ECON_ varieerusid vaid kuni 5%, kuid kontrastsusega kaalutud servade tihedus CWED varieerus 27-55% (joonis 23). ECON_ ja TECI arvutatakse protsentides ehk neil indeksitel on kindel ulatus, kuid indeksil CWED puudub väärtuse ülempiir ja ühikuks on meetrit hektari kohta, mistõttu CWED sõltub servapikkusest. Servapikkus omakorda sõltub eraldiste

arvust, mida eCognition-iga klassifitseeritud Tartu kogu ala kujutaval pildil oli 58% rohkem kui Idrisiga klassifitseeritud pildil.

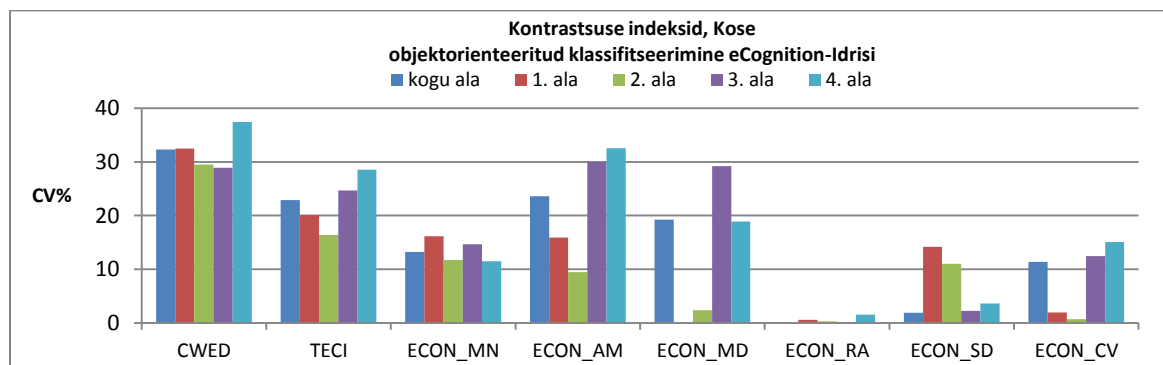
Kose alade objektorienteeritud klassifitseerimistulemuste põhjal arvutatud kontrastsuse indeksid varieerusid rohkem kui teiste töötlusviiside ja alade lõikes (joonis 24). Põhjuseks on asulate ja infrastruktuuri klass, mis eCognition-iga klassifitseerides eristus paremini ning selle osakaal maastikus oli suurem. Kuna asulate/infrastruktuuri klassile määrati kontrastsuse maatriksiga suur kontrastsuse kaal teiste klasside suhtes, siis varieerusid Idrisi ja eCognition-i tulemused selle klassi erineva osakaalu tõttu maastikus üsna palju.

Tabel 20. Kontrastsuse indeksite keskmised variatsioonikoefitsendid töötlusviiside kaupa järjestatud kasvavalt kõikide alade keskmise järgi.

	Kose					Tartu				
	ISODATA Erdas	ISODATA Erdas vs Idrisi	k-means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	k-means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	kõik alad keskmise	kõik alad mediaan
ECON_RA	1.31	0.47	0.06	0.42	0.50	3.71	3.83	0.61	1.36	0.53
ECON_MN	8.41	0.87	1.42	0.80	13.46	1.06	0.53	1.99	3.57	1.44
ECON_MD	14.76	0.37	1.41	1.81	13.94	0.00	1.42	1.73	4.43	0.00
ECON_CV	8.27	2.77	5.65	0.73	8.32	3.80	4.09	2.68	4.54	3.94
ECON_SD	10.47	3.13	6.52	0.49	6.59	3.34	4.62	2.22	4.67	3.65
TECI	9.33	1.73	5.89	0.55	22.52	0.99	0.40	2.54	5.49	2.13
ECON_AM	10.78	2.24	7.21	0.57	22.30	1.31	0.45	4.92	6.22	2.35
CWED	13.04	2.00	10.40	0.69	32.13	9.05	4.55	41.50	14.17	8.27



Joonis 23. Kontrastsuse indeksite variatsioonikoefitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Tartu alal alade kaupa.



Joonis 24. Kontrastsuse indeksite variatsioonikoefitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Kose alal alade kaupa.

Mitmekesisuse indeksid

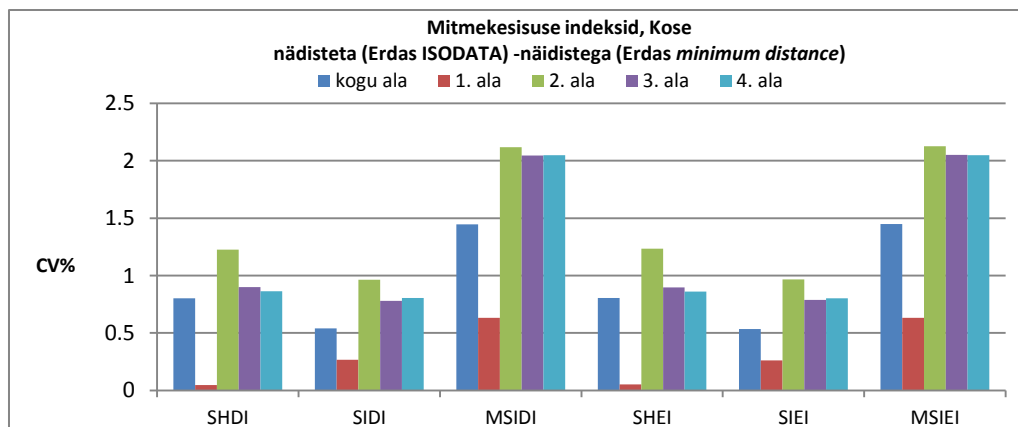
Kõige vähem varieerusid kõikidest indeksite tüüpidest mitmekesisuse indeksid, mille variatsioonikoeffitsendid jäid alla 20% ja paljudel juhtudel varieerusid vaid mõne protsendi ulatuses (joonis 25). Kõikide alade keskmised variatsioonikoeffitsendid jäid alla 5% (tabel 21). Mitmekesisuse indeksid sõltuvad kahest peamisest komponendist: rohkusest ja ühtlusest (McGarigal *et al.* 2012). Kuna rohkus ehk klasside / kategooriate arv oli juba eelnevalt määratud ning kõigil võrdluspaaridel sama, siis erines vaid klasside / kategooriate osakaal maastikus. Kõikide alade ja töölusviiside puhul varieerusid kõige vähem Simpsoni mitmekesisuse (SIDI) ning Simpsoni ühtluse (SIEI) indeksid ja kõige rohkem muudetud Simpsoni mitmekesisuse (MSIDI) ning muudetud Simpsoni ühtluse (MSIEI) indeksid, kuigi erinevused ei olnud suured.

Objektorienteeritud klassifitseerimisel varieerusid teistest veidi enam nii Tartu kui Kose 3. ja 4. ala indeksid (joonis 26-27). Kose 3. ja 4. alal olid asustuse eraldised, mis eCognition tarkvarga klassifitseerides tulid paremini esile ja olid suuremad kui Idrisi-ga. Lisaks olid Kose 3. ja 4. ala heterogeensema maakasutusega (erineva spektraalse tundlikkusega põllud, asustus, mets, infrastruktuur), mis suurendas klassifitseerimise erinevusi. Tartu 3. ja 4. alal oli palju metsa, millega kaasnesid loodesuunalised varjud, mis eristusid hästi eCognition tarkvaraga klassifitseerides, kuid halvasti Idrisi-ga klassifitseerides.

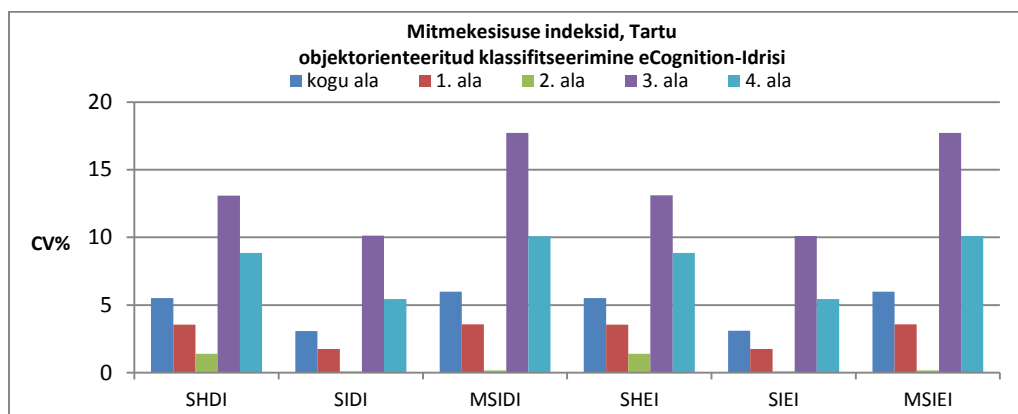
Mitmekesisuse indeksite puhul eristusid selgelt üksteisega täielikult korreleeruvad paarid, mille variatsioonikoeffitsendid olid alade lõikes ühesugused: muudetud Simpsoni mitmekesisuse (MSIDI) ja muudetud Simpsoni ühtluse (MSIEI) indeksid; Simpsoni mitmekesisuse (SIDI) ja Simpsoni ühtluse (SIEI) indeksid; Shannoni mitmekesisuse (SHDI) ja Shannoni ühtluse (SHEI) indeksid. Enamiku mitmekesisuse indeksite väärtused arvutatakse vahemikus 0 kuni 1, erandiks on SHDI ja MSIDI, mille vahemik on 0 kuni lõpmatus, kuid käesolevas töös jäid ka nende indeksite väärtused kindlasse vahemikku 1 kuni 2.

Tabel 21. Mitmekesisuse indeksite keskmised variatsioonikoeffitsendid töölusviiside kaupa järjestatud kasvavalt kõikide alade keskmise järgi.

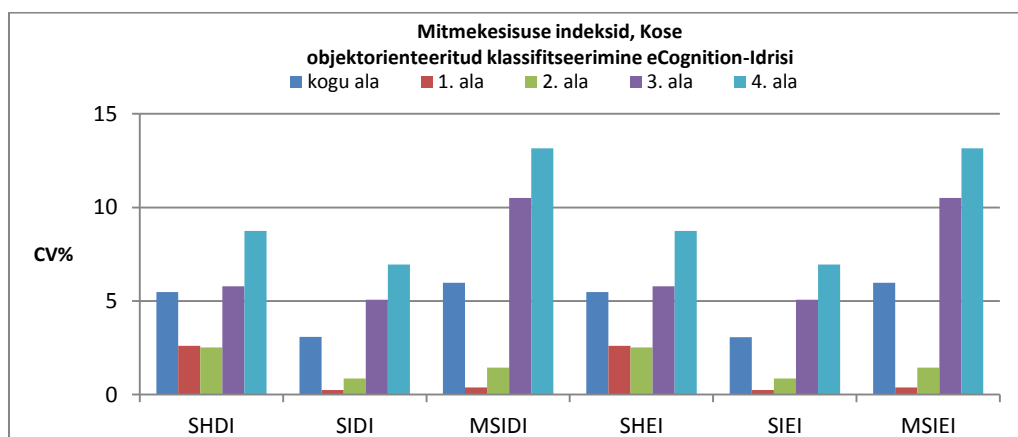
	Kose					Tartu				
	ISODATA Erdas	ISODATA Erdas vs Idrisi	k-means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	k-means ENVI vs Idrisi	nädisteta vs nädistega	obj. klassif. eCognition vs Idrisi	kõik alad keskmise	kõik alad mediaan
SIDI	3.27	0.86	0.92	0.67	3.24	4.42	1.33	4.10	2.35	1.20
SIEI	3.27	0.86	0.92	0.67	3.24	4.42	1.33	4.10	2.35	1.20
SHEI	5.01	1.11	1.39	0.77	5.03	6.31	2.59	6.48	3.59	2.51
SHDI	5.02	1.11	1.39	0.77	5.03	6.31	2.60	6.48	3.59	2.52
MSIDI	7.81	2.03	2.19	1.66	6.29	8.93	2.80	7.51	4.90	2.49
MSIEI	7.81	2.03	2.19	1.66	6.29	8.94	2.80	7.51	4.90	2.49



Joonis 25. Mitmekesisuse indeksite variatsioonikoeffitsendid (CV%) näidistega ja näidisteta klassifitseerimisel Kose alal alade kaupa.



Joonis 26. Mitmekesisuse indeksite variatsioonikoeffitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Tartu alal alade kaupa.



Joonis 27. Mitmekesisuse indeksite variatsioonikoeffitsendid (CV%) tarkvaradega eCognition ja Idrisi objektorienteeritud klassifitseerimisel Kose alal alade kaupa.

4. Arutelu

4.1. Tarkvaratöötlus

Näidisteta klassifitseerimise ehk klasterdamise algoritmide võrdlemisel tarkvarades Erdas, ENVI, Idrisi selgus, et samade või sarnaste parameetrite korral võivad tulemused varieeruda üsna palju. Idrisi ISODATA ja k-means tulemused erinesid ENVI ja Erdas-e omadest kohati üsna palju, kusjuures samade parameetritega, sh sama klastrite algamise reeglina korduval katsetamisel võisid tulemused tulla Idrisi puhul erinevad / juhuslikud. Samade parameetritega korduval katsetamisel andsid ENVI ja Erdas-e klasterdamise algoritmid aga sama klassifitseerimistulemuse. Peamine varieeruvust põhjustav parameeter oli klastrite algamise reegel, mis määrab esialgsetele klastritele keskmised heledusväärtused. Ka erinevad autorid tõdevad, et klassifitseerimise lõpptulemus sõltub paljuski pikslite esialgsest jaotumisest klastrite vahel (Peña *et al.* 1999; Huang 2002; Redmond, Heneghan 2007, Jain 2010). K-means algoritmi eesmärk pikslite klastritesse määramisel on vähendada klastrisisest varieeruvust, heledusväärtuste ruutkauguste (vigade) summa peab olema minimaalne iga piksli ja klatri keskpunkti vahel. Teisalt läheneb k-means seejuures vaid lokaalse miinimumini, mistõttu parima tulemuse saavutamiseks tuleks algoritmi tööd katsetada mitmeid kordi (Peña *et al.* 1999; Jain 2010). ENVI ja Erdas-e klassifitseerimistulemuste ühtivuse ning korratavuse võimalikuks põhjuseks on ühesugune klastrite algamise reegel – *diagonal axis*, mille puhul määratakse klastrite keskpunktid alati ühte moodi – heledusväärtuste diagonaalsel teljel võrdsete vahemike järel n-dimensionaalses ruumis (n – spektraalkanalite arv). Ka Erdas ISODATA teine klastrite algamise reegel – *principal axis*, andis korduval katsetamisel samad tulemused, siingi määratakse klastrite algsed keskpunktid alati ühtemoodi – peakomponentide teljel. Teisalt on kummaline, et Idrisi Selva klastrite algamise reeglit *diagonal axis* kasutades varieerusid tulemused üsna palju. Idrisi klastrite algamise reeglitega juhuslik jagunemine (*random partition*) ja juhuslikud keskpunktid (*random seed*, kirjanduse järgi ka *Forgy*) võib pikslite esialgne jaotumine olla iga kord erinev, mistõttu ka tulemused varieeruda. Teisalt peaksid need olema standard klastrite algamise reeglid k-means algoritmi jaoks, kuigi ka mitmeid teisi on loodud (Peña *et al.* 1999; Redmond, Heneghan 2007).

Erdas-e ja ENVI ISODATA algoritmid ühtisid Erdas-e diagonaalse telje (*diagonaal axis*) algamise reeglit kasutades täielikult, millest võib järeldada, et ka ENVI algoritm kasutab seda sama reeglit. Erdas-e ISODATA algoritm nõuab klastrite määramisel täpset arvu, kuigi algse definitsiooni järgi peaks kasutaja saama määrata võimaliku vahemiku (Ball, Hall 1965). Seetõttu on Erdas-e ISODATA algoritm pigem k-means algoritm. Teisalt ei ole see nimetus ka väga vale, sest ISODATA ongi tegelikult k-means algoritmi modifikatsioon (Jain 2010). Ka ENVI k-means ja Erdas-e ISODATA klasterdamise algoritmid andsid samaväärsed klassifitseerimistulemused loodavate klastrite arvu ühesel määramisel, mis kinnitab, et ISODATA on tõepoolest algoritmi k-means modifikatsioon.

Kose alal erinesid Erdas ja Envi ISODATA klassifitseerimistulemused 0,03%, mis võib olla tingitud parameetrist iteratsiooni käigus klassi mitte muutvate pikslite protsent, mis käesolevas töös oli 95%. Seega 5% pikslitest võisid jääda „ebasobivasse“ klassi ning järgnevate iteratsioonidega oleksid võinud veel oma klastreid muuta. Parameetri väärtuse korral 100% oleks aga algoritmi töö muutunud tunduvalt aeganõudvamaks, kuid tulemusi oleks see väga vähe mõjutanud, mida kinnitab ka vaid 0,03% erinevus.

Etalonide ehk õpetuspiirkondadega klassifitseerides ühtisid eri tarkvarade samanimeliste algoritmide klassifitseerimistulemused see-eest väga hästi. Nii lühima tee (*Minimum Distance*) kui suurima tõepära (*Maximum Likelihood*) algoritmiga klassifitseerides ei erinenud kummalgi testalal tulemused üle 5%. Seega ei ole vahet, millist tarkvara neist kolmest kasutada näidistega klassifitseerimisel, kui õpetuspiirkonnad on samad. Teiste õpetuspiirkondadega oleksid tulnud ilmselt mõnevõrra teistsugused klassifitseerimistulemused, kuigi tõenäoliselt sarnased, kuna õpetuspiirkonnad valiti näidisteta klassifitseerimise põhjal. Näidisteta klassifitseerimist saabki kasutada sobivate õpetuspiirkondade puudumisel andmekaevena, kuna klasterdamine ei eelda palju eelteadmisi ning on odavam kui välitööd (Jain *et al.* 1999). Teisalt võib klasterdamise tulemusi olla keeruline tõlgendada, kui sarnase spektraalväärtusega klassid ei eristu üksteisest hästi.

GeoEye pildi etalonideta klassifitseerimisel tekitas probleeme klastritele nimetuste määramine, kuna maakatteklasside spektraalväärtused olid üsna homogeensed ega eristunud üksteisest hästi. Näiteks segunesid omavahel metsa ja põllumaade klastrid. Nii GeoEye kui RapidEye pildi puhul ei olnud võimalik klasterdamisega teistest klassidest eristada hoonete ja infrastruktuuri maakatte tüüpe, mille spektraalsed väärtused sarnanesid põldude omadega. Lisaks on suure ruumilise lahutusega satelliidipiltidel asustuse pikslid üsna heterogeensed, mistõttu on neid raske paigutada ühte klastrisse. Suure ruumilise lahutusega piltide klassifitseerimisel on seetõttu parem kasutada objektorienteeritud klassifitseerimist, võrreldes pikselhaaval klassifitseerimisega (Lu, Weng 2007). See aitab vältida mürapikslite teket ehk üksikute ümbruskonnast erinevate pikslite klassifitseerimist eraldi klassidesse (ingl *salt and pepper effect*). Mürapikslite probleem tekkis eriti Tartu pildi puhul, mille ruumiline lahutus oli väga suur (2 m). Objektorienteeritud klassifitseerimine võtab arvesse ümbritsevaid piksleid ja pikslite ruumilist paiknemist. Keerukamate maastike / maastikuklasside korral võib maakatte seegmine spektraalne varieeruvus olla väga suur, mistõttu oleks tarvis klassifitseerimisel arvestada ka maastiku tekstuuri (Lu, Weng 2007). Käesolevas töös olid sellisteks probleemseteks maastikuklassideks mets ja asustus, mis pikselhaaval klassifitseerides ei eristunud hästi, seda eriti Tartu alal. Etalonidega klassifitseerimisel oleks ehk abi olnud metsa ja asustuse mittepameetrite signatuuride loomisest, mis sobivad spektraalselt kattuvate ning normaaljaotusele mitte vastavate alade eristamiseks (Cortijo, Perez de la Blanca 1997), kuid kuna eesmärgiks oli võrrelda ka etalonidega klassifitseerimist etalonideta klassifitseerimisega, siis oleks see suurendanud erinevusi nende kahe klassifitseerimisviisi vahel. Mürapikslite eemaldamiseks saab kasutada ka filtreid (Alamri *et al.* 2010). Käesolevas töös kasutati Tartu alal etalonidega ja etalonideta klassifitseerimisel selleks sobivat 5x5 aknaga mediaanfiltrit. Klassifitseerimistäpsus sellest ei muutunud.

Indeksite arvutamiseks ja võrdlemiseks kasutatud eri tarkvarade, kuid samanimeliste või sarnaste algoritmide ning töötlusviiside klassifitseerimistulemused ühtisid käesolevas töös 75-90%. Klasside lõikes olid erinevused suuremad, ulatudes kohati isegi 50%-ni. Klasside erinevused mõjutavad omakorda ruumilisi erinevusi. Ka erinevad autorid rõhutavad, et kaartidel võib olla hea klassifitseerimistäpsus, kuid see ei ütle veel midagi vigade klassidevahelise või ruumilise jagunemise kohta, mis võib oluliselt mõjutada maastikumustri kirjeldamise täpsust (Steele *et al.* 1998; Shao *et al.* 2001; Fassnacht *et al.* 2006; Langford *et al.* 2006). See seletab ilmselt ka klassi taseme indeksite suuremat tundlikkust klassifitseerimistäpsuse suhtes, mida teised autorid on leidnud (Mas *et al.* 2010).

Lisaks kaugseire pilditööstlustarkvarade töötlustehnikatele ja klassifitseerimisviisidele, sõltuvad indeksite väärtused ka maastikumeetrika arvutamiseks mõeldud tarkvara parameetritest. Funktsionaalsete indeksite puhul peab kasutaja defineerima näiteks servasügavuse, arvutamise kaugused, sarnasuse ning servakontrastsuse maatriksid ning selle, kas arvutustes kasutatakse 4- või 8-naabri reeglit.

4.2. Maastikuindeksid

Indeksite tüüpidest olid kõige stabiilsemad mitmekesisuse indeksid ning kõige enam varieerusid tuumala indeksid ning mõned koonduvuse indeksid (tükeldatuse SPLIT, koondatuse MESH, sarnasuse SIMI ning läheduse PROX indeksid), mis ühtib Mas *et al.* (2010) tulemustega. Võrreldes indeksite kõikide alade ja töötluste keskmisi variatsioonikoefitsiente mediaan variatsioonikoefitsientidega (lisa 3), siis peaaegu kõigil juhtudel olid mediaan väärtused kas samas suurusjärgus või väiksemad. Seega suurendasid enamasti üksikud võrdluspaarid keskmisi varieeruvusi.

Mitmekesisuse indeksitest omakorda varieerusid kõige vähem omavahel korreleeruvad Simpsoni mitmekesisuse indeks SIDI ja Simpsoni ühtluse indeks SIEI, mis on vähem tundlikud harva esinevate eraldiste tüüpide suhtes kui Shannoni samanimelised indeksid (McGarigal *et al.* 2012). Käesolevasse töösse ei kaasatud eraldiste rohkust kirjeldavaid mitmekesisuse indekseid, kuna need põhinevad eritüübiliste eraldiste arvul, mis oli autori poolt määratud ja püsis muutumatu kõigil testaaladel. Viimane võib olla üheks põhjuseks, miks mitmekesisuse indeksid, mis mõõdavad maastiku kompositsiooni ning põhinevad eraldiste rohkusel ja eraldiste kategooriate osakaalul maastikus (McGarigal, Marks 1995), varieerusid käesolevas töös väga vähe. Kui ISODATA algoritmiga klassifitseerimisel oleks kasutatud loodavate klastrite arvu määramisel vahemikku (võimalik Idrisi ja ENVI ISODATA moodulites), siis oleksid mitmekesisuse indeksid rohkem varieerunud, sest eri takvarades klassifitseeriti sama pildi, algoritmi ja klastrite vahemiku korral erinev arv klastreid. Teisalt oleks see muutnud indeksite varieeruvuste põhjendamise keerulisemaks, sest suur osa indekseid, nii kompositsiooni kui konfiguratsiooni kirjeldavaid, on tundlikud muutustele temaatilises resolutsioonis (Baldwin *et al.* 2004; Buyantuyev, Wu 2007; Peng *et al.* 2007).

Kuju indeksitest varieerusid teistest vähem fraktaalset dimensiooni kirjeldavad indeksid (FRAC_, PAFRAC), mis põhinevad übermõõdu ja pindala logaritmilistel arvutustel, sõltudes seega vähem eraldiste suurusest, erinevalt SHAPE ja PARA indeksitest. Stabiilne oli ka väikseima võimaliku ümbritseva ringi indeksite rühm CIRCLE_, mis peaks sobima piklike ja välja venitatud eraldiste kirjeldamiseks maastikul (McGarigal *et al.* 2012). See indeksite rühm ei varieerunud isegi eriti Tartu objektorienteeritud klassifitseerimiste puhul, kus eCognition tarkvaraga eristusid piklikud varjud palju paremini kui Idrisi-ga.

Tuumala indeksite hulgas peaaegu ei olnudki väga stabiilseid indekseid, mis võib olla tingitud sellest, et neid indekseid mõjutavad mitmed tegurid: kasutaja määratud servasügavus, eraldiste kuju ja suurus (McGarigal, Marks 1995). Kuigi kahe indeksi, tuumalade pindalade mediaan CORE_MD ja tuumalade osakaalude mediaan eraldistes CAI_MD, variatsioonikoefitsendid olid 0-id, siis on see tingitud ilmselt paljude väikeste või piklike eraldiste suurest osakaalust maastikus, seda isegi objektorienteeritud klassifitseerimiste korral, kus pindalad olid suuremad. Tuumala indeksid ei erista suuri ja keeruka kujuga eraldisi väikestest eraldistest – mõlema tuumala võib olla väike, sõltumata pindalast.

Koonduvuse indeksitest olid väga ebastabiilsed sarnasuse (SIMI_) ja läheduse PROX_ indeksid ning üksteisega täielikult korreleeruvad MESH (koondatus) ja SPLIT (tükeldatus), mis kõik sõltuvad käesolevas töös palju varieerunud pindala indeksitest. Mas *et al.* (2010) tõdeb, et eraldatuse / läheduse hindamiseks tasub pigem kasutada lähima naabri indeksid ENN_, võrreldes PROX_ indeksitega, ning koonduvuse / vahelduvuse hindamiseks AI, PLADJ, DIVISION, CONTAG indeksid, mis ka käesolevas uurimuses olid stabiilsemad kui indeksid MESH ja SPLIT.

Stabiilsete indeksite hulgas oli enamikul indeksitest kindel väärtuste vahemik ehk ülem- ja alampiir. Üldiselt on ka lihtsam interpreteerida indeksid, mille väärtus jääb kindlasse vahemikku (Li, Wu 2004).

Muutusi fraktaalses dimensioonis ja koonduvuses on keeruline tõlgendada, kuna mitmed tegurid mõjutavad korraga indeksi väärtust ja need ei moodustu lineaarselt (Li, Wu 2004). Käesolevas töös olid need indeksid stabiilsed, see-eest lihtsasti tõlgendatavad indeksid, nagu enamik pindala ja serva indeksitest (eraldiste suurus AREA, suurima eraldise osakaal LPI), varieerusid üsna palju, sõltuvalt klassifikatsioonist ja alast. See ühtib Shao, Wu (2008) hinnangutega, et enamasti on maastiku struktuuri hästi kirjeldavad indeksid tundlikumad klassifitseerimisvigade või erinevuste suhtes. Käesolevas töös oli siiski ka selliseid kergemini interpreteeritavaid indeksid, mille variatsioonikoefitsiendid ei olnud üldiselt väga suured: lähima naabri indeksid ENN_, servatihedus ED, eraldiste arv NP või eraldiste tihedus PD (v.a objektorienteeritud klassifitseerimisel). Ka Lechner *et al.* (2013) tulemustest selgub, et satelliidipiltide töötamise ja tegeliku maastikumustri mõju indeksitele on raske üksteisest lahus hoida. Näiteks NP oli selle uurimuse järgi tundlik nii töötamise kui maastikumustri suhtes.

Hinnates eri töötlusviiside ja üldise klassifitseerimistäpsuse mõju indeksite varieeruvusele, ei saa kindlalt väita, et mida suurem klassifitseerimistäpsus, seda stabiilsemad indeksid, nagu teised autorid on leidnud (Shao *et al.* 2001; Altamirano *et al.* 2012). Näiteks ei varieerunud väga palju k-means algoritmiga ENVI ja Idrisi tarkvarades klassifitseeritud Kose ala piltidelt arvutatud indeksid, kuigi piltide klassifikatsioonide ühtivus oli käesoleva töö kontekstis teiste pildipaaridega võrreldes madal – 76%. Üldiselt varieerusid rohkem objektorienteeritud klassifitseerimistega saadud indeksid, võrreldes pikselhaaval klassifitseeritud piltidelt arvutatud indeksitega. Nende klassifitseerimiste ühtivus oli ka madal, 75-79%. See-eest ISODATA algoritmiga Erdas ja Idrisi tarkvarades klassifitseeritud Kose ala piltidelt, mille klassifikatsioonide ühtivus oli 90%, arvutatud indeksid varieerusid kõige vähem. Shao *et al.* (2001) järgi saavutavadki enamik maastikuindekseid stabiilsed väärtused klassifitseerimistäpsuse juures, mis on üle 90%. Objektorienteeritud Kose ja Tartu pildi puhul suurendas varieeruvust osade maakatteklasside (infrastruktuur, asustus, varjud) halb eristumine Idrisi tarkvaraga, mis mõjutas maastiku keerukust, fragmenteeritust, servi. McGarigal, Marks (1995) tõdeb, et maastikuindeksite arvutamise puhul on oluline servade definitsioon: suure osa indeksite väärtus võib kõvasti muutuda, sõltuvalt servade keerukusest.

Tartu GeoEye satelliidipilt oli suurema ruumilise lahutusega kui Kose pilt ning pildistatud oktoobris, mil varjud on pikemad kui suvel, mistõttu eristusid need omaette klassina. Pikselhaaval klassifitseerides ühtisid ENVI ja Idrisi k-means algoritmiga klassifitseeritud varjude klassid kahel pildil hästi, kuid objektorienteeritud klassifitseermisel ei eristunud need Idrisi-ga nii hästi, mis suurendas mitmete indeksite, nagu servapikkus või mõned kuju kirjeldavad indeksid, varieeruvust.

Kuna indekseid on palju rohkem kui parameetreid, millel need põhinevad, siis paljud indeksid korreleeruvad (McGarigal, Marks 1995), mis ilmnes ka variatsioonikoeffitsientides. Identsed variatsioonikoeffitsendid andsid järgmised indeksite paarid: TE – ED, NP – PD, DCAD – NDCA, MSIDI – MSIEI, SHDI – SHEI, SIDI-SIEI, MESH – SPLIT – AREA_AM, mistõttu neid indekseid ei ole üldiselt mõtet koos kasutada.

Kui võrrelda Leitão *et al.* (2006) poolt välja pakutud tuumik indeksite loendit, mille indeksid peaksid kirjeldama kõige paremini maastiku eri aspekte, käesoleva töö tulemustega, siis selgub, et pooled neist (GYRATE, CONTAG, ECON, ENN) peaksid olema erinevate klassifitseerimisviiside suhtes stabiilsed ning NP / PD olid keskmise varieeruvusega, pooled (AREA_AM, AREA_MN, SHAPE ja PROX) aga ebastabiilsed, kuigi kuju indeksi SHAPE statistilist jaotust kirjeldavad variandid SHAPE_MN ja SHAPE_MD olid samuti stabiilsed. Uuemaa *et al.* (2011) uurisid peakomponentide analüüsiga indeksite korreleerumist Eesti maastikel, et samuti leida tuumik indekseid. Neli olulisemat maastiku ruumilise struktuuri aspekte kirjeldavat indeksit olid servatihedus ED, kogu serva kontrastsus TECI, keskmine kuju indeks SHAPE_MN, eritüübiliste eraldiste tihedus PRD. SHAPE_MN, nagu juba mainitud, oli stabiilne, TECI ja ED olid keskmise stabiilsusega, varieerudes palju vaid kas Kose või Tartu alade objektorienteeritud klassifikatsioonide korral, PRD jäi arvutustest välja, kuna eritüübiliste eraldiste arv püsis kõigi võrdluspaaride lõikes muutumatu. Mas *et al.* (2010), leiab, et eraldiste rohkust kirjeldavad indeksid võivad olla liigagi stabiilsed, sh maastiku muutuste suhtes.

Kokkuvõte

Töö peamine eesmärk oli uurida, millised maastikuindeksid on satelliidipiltide töötamise suhtes kõige stabiilsemad ning kuidas erinevad kommertstarkvarade algoritmid ning klassifitseerimisviisid neid mõjutavad.

Nelja tarkvara klassifitseerimismeetodite võrdlemisel Kose ja Tartu testalade peal selgus, et samanimeliste algoritmidega võivad üsna erinevad klassifitseerimistulemused tulla. Etalonideta klassifitseerimisel ühtisid ISODATA algoritmiga saadud tulemused tarkvarades Erdas ja ENVI täielikult, kuid Idrisi Selva omad erinesid neist 2-38%, ning k-means algoritmi klassifitseerimistulemused ühtisid ENVI ja Idrisi tarkvarades 64-86%. Peamine varieeruvust põhjustav parameeter oli klastrite algamise reegel, mis määrab esialgsetele klastritele keskmised heledusväärtused ning pikslite esialgse jaotumise klastrite vahel. Kuna Idrisi algoritmide ISODATA ja k-means klastrite algamise reeglid määravad esialgsed klastrid juhuslikult, erinevalt Erdas-e ja ENVI omadest, siis varieerusid Idrisi etalonideta klassifitseerimistulemused üsna suures ulatuses, kuid Erdas ja ENVI omad olid samade parameetrite korral stabiilsed.

Näidistega klassifitseerimisel erinesid Erdas, ENVI ja Idrisi lühima tee (*Minimum Distance*) ja suurima tõepära (*Maximum Likelihood*) algoritmide klassifikatsioonid vastavalt alla 1% ja alla 5%, mistõttu ei ole suurt vahet, millist tarkvara neist kolmest kasutada nende algoritmidega klassifitseerimisel, kui õpetuspiirkonnad on samad. Kuna õpetuspiirkonnad valiti Erdas ISODATA klassifitseerimistulemuste põhjal, siis võrreldi näidisteta klassifitseerimist ka näidistega klassifitseerimisega (algoritmi *Minimum Distance*). Mõlemal alal ühtisid tulemused 89%.

Suure ruumilise lahutusega piltide klassifitseerimiseks sobib paremini objektorienteeritud klassifitseerimine, mis arvestab maastiku tekstuuri ning aitab vältida mürapikslite teket. Objektorienteeritud klassifitseerimise võimalusi võrreldi tarkvarades eCognition ja Idrisi Selva. Tulemused ühtisid Kose alal 79% ja Tartu alal 75%. Klasside lõikes olid nii objektorienteeritud kui ka teiste võrreldud klassifitseerimisviiside erinevused suuremad, ulatudes kohati isegi 50%-ni.

Maastikuindeksite tüüpidest olid kõige stabiilsemad mitmekesisuse indeksid ning kõige ebastabiilsemad tuumala indeksid, kuigi ka teiste indeksite tüüpide hulgas leidis indekseid, mille variatsioonikoefitsiendid olid kõrged. 111-st arvutatud maastiku taseme indeksist olid tarkvara töötamise suhtes stabiilsed umbes kolmandik, teine kolmandik varieerusid suures ulatuses ning ülejäänud olid keskmise stabiilsusega, varieerudes üle 10% vaid üksikute alade või töötluste lõikes. Lisaks tuumala indeksitele varieerusid suurel määral mõned koonduvuse indeksid: tükeldatuse SPLIT, koondatuse MESH, sarnasuse SIMI, läheduse PROX ning pindala AREA indeksid. Stabiilsete indeksite hulgas oli enamikul indeksitest kindel väärtuste vahemik ehk ülem- ja alampiir.

Üldiselt varieerusid rohkem objektorienteeritud klassifitseerimiste põhjal arvutatud indeksid, võrreldes pikselhaaval klassifitseeritud piltide omadega. Põhjuseks on osade maakatteklasside (infrastruktuur, asustus, varjud) halb eristumine Idrisi tarkvaraga klassifitseerimisel, võrreldes spetsiaalselt objektorienteeritud klassifitseerimise tarvis loodud eCognition tarkvaraga.

Võrreldes käesoleva töö tulemusi teiste autorite (Leitão *et al.* 2006; Uuemaa *et al.* 2011) välja pakutud soovituslike tuumik indeksitega, mis peaksid kirjeldama kõige paremini maastiku eri aspekte, siis selgub, et üle poolte neist on stabiilsed (GYRATE, CONTAG, ECON, ENN, SHAPE_MN) või enamasti stabiilsed (NP/PD, ED, TECI).

Summary

Sensitivity of landscape pattern metrics to processing methods of satellite images

Landscape metrics i.e landscape indices quantify landscape structure or pattern. Spatial structure influences ecological processes and has a close relationship with biodiversity (McGarigal & Marks 1995). In most cases landscape metrics are derived from satellite images and image processing software is used prior calculating landscape metrics. However, different properties of satellite images and geospatial processing methods can cause spatial uncertainty which affects the representation of landscape pattern. For example, pixel size i.e spatial resolution (Saura 2004; Wu 2004; Símová, Gdulová 2012), classification scheme and thematic resolution (Baldwin *et al.* 2004; Buyantuyev, Wu 2007; Peng *et al.* 2007) and poor classification accuracy (Shao *et al.* 2001; Langford *et al.* 2006; Altamirano *et al.* 2012) are the main factors according to literature causing misinterpretation of landscape pattern calculated from satellite images. Nevertheless, few studies have addressed the uncertainty of landscape metrics caused by different image processing methods or software.

The main purpose of this study was to determine which landscape indices are the least sensitive to different kinds of image processing methods and classification algorithms used in remote sensing. Four commercial software packages were used: Erdas Imagine 2011, ENVI 5.0, Idrisi Selva, eCognition 8.9. Processing methods that were compared in these applications were unsupervised classification (algorithms ISODATA and k-means), supervised classification (algorithms Minimum Distance and Maximum Likelihood) and object-oriented classification (eCognition and Idrisi Segmentation Classifiers). Two study areas were classified by using these different approaches.

Although of the same name, clustering algorithms ISODATA and k-means differ in different image processing softwares and classifications made by these algorithms varied to great extent. The main reason for that is the initialising method which defines the initial clusters and affects the final results. Supervised classifications using algorithms Minimum Distance and Maximum Likelihood did not differ very much among Erdas, ENVI and Idrisi but comparing supervised classifications, which training areas were derived from unsupervised classification (ISODATA), with unsupervised classifications the agreement was 89%. Object-oriented classifications made in eCognition and Idrisi differed up to 25%, although variations were greater among different landcover classes.

Landscape level metrics were calculated in spatial pattern analysis program Fragstats from 8 pairs of classified images using above mentioned processing methods in order to evaluate their sensitivity to these classifications. The coefficients of variation were calculated for each pair of indices. The most stable group of landscape level metrics were the diversity metrics which indices had low coefficients of variation, and the most sensitive to different kinds of classification methods were the core area metrics. From 111 calculated metrics approximately one third were stable indices, one third were unstable and the rest had mostly low variation levels with few exceptions. Besides core area metrics some aggregation and area metrics showed high coefficients of variation, these were the Splitting index SPLIT, the Effective Mesh size MESH, the similarity indices SIMI, the proximity indices PROX and the area indices AREA. Most of the stable indices in this study had specific range of values i.e minimum and maximum values which also makes them easier to interpret.

Landscape metrics calculated from object-oriented classified satellite images had tendency to vary more than those calculated by pixel-based classification because Idrisi could not differentiate some of the landcover classes (infrastructure, buildings, shades) from others comparing to eCognition.

Kasutatud kirjandus

- Al-amri, S., S., Kalyankar, N., V., Khamitkar, S., D., 2010.** A Comparative Study of Removal Noise from Remote Sensing Image. *International Journal of Computer Science Issues* 7(1), 32-36.
- Altamirano, A., Miranda, A., Jiménez, C., 2012.** Uncertainty of landscape indices for spatial structure analysis. *Bosque* 33(2), 171-181.
- Anderson, J., R., Hardy, E., Roach, J., Witmer, R., 1976.** A land use and land cover classification system for use with remote sensor data. US Government Printing Office, Washington.
- Baldwin, D., J., B., Weaver, K., Schnekenburger, F., Perera, A., H., 2004.** Sensitivity of landscape pattern indices to input data characteristics on real landscapes: implications for their use in natural disturbance emulation. *Landscape Ecology* 19(3), 255-271.
- Ball, G., H., Hall, D., J., 1965.** ISODATA. A novel method of data analysis and pattern classification. Stanford Research Institute, Menlo Park, United States. Office of Naval Research. Information Sciences Branch.
- Brown, D. G., Addink, E., A., Duh, J., D., Bowersox, M., A., 2004.** Assessing uncertainty in spatial landscape metrics derived from remote sensing data. In R. S. Lunetta and J. G. Lyon (eds), *Remote Sensing and GIS Accuracy Assessment*. CRC Press, Boca Raton. Chapter 16: 221-232.
- Buyantuyev, A., Wu, J., 2007.** Effects of thematic resolution on landscape pattern analysis. *Landscape Ecology* 22, 7-13.
- Cortijo, F., J., Perez de la Blanca, N., 1997.** A comparative study of some non-parametric spectral classifiers. Applications to problems with high-overlapping training sets. *International Journal of Remote Sensing* 18(6), 1259-1275.
- Dramstad, W., E., 2009.** Spatial metrics – useful indicators for society or mainly fun tools for landscape ecologists? *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography* 63(4), 246-254.
- eCognition Developer 8.9 User Guide 2013.** Trimble Germany GmbH, Munich.
- Ediriwickrema, J., Khorram, S., 1997.** Hierarchical Maximum-Likelihood Classification for Improved Accuracies. *IEEE Transactions on Geoscience and Remote Sensing* 35(4), 810-816.
- Fassnacht, K., S., Cohen, W., B., Spies, T., A., 2006.** Key issues in making and using satellite-based maps in ecology: A primer. *Forest Ecology and Management* 222(1-3), 167-181.
- Foody, G., M., 2008.** Harshness in image classification accuracy assessment. *International Journal of Remote Sensing* 29(11), 3137-3158.

- Huang, K., 2002.** A Synergistic Automatic Clustering Technique (SYNERACT) for Multispectral Image Analysis. *Photogrammetric Engineering & Remote Sensing* 68(1), 33-40.
- Jain, A., K., Murty, M., N., Flynn, P., J., 1999.** Data Clustering: A Review. *ACM Computing Surveys* 31(3), 264-323.
- Jain, A., K., 2010.** Data clustering: 50 years beyond K-means. *Pattern Recognition Letters* 31(8), 651-666.
- Jensen, J., R., Im, J., Hardin, P., Jensen, R., R., 2009.** Image classification. *Remote Sensing of the Environment*, T. A., Nellis, M., D., Foody, G., M. (eds.), The Sage Handbook of Remote Sensing. SAGE, London, 269-281.
- Kelly, M., Tuxen, K., A., Stralberg, D., 2010.** Mapping changes to vegetation pattern in a restoring wetland: Finding pattern metrics that are consistent across spatial scale and time. *Ecological Indicators* 11, 263-273.
- Langford W., T., Gergel, S., E., Dietterich, T., G., Cohen, W., 2006.** Map misclassification can cause large errors in landscape pattern indices: Examples from habitat fragmentation. *Ecosystems* 9(3), 474-488.
- Lausch, A., Herzog, F., 2002.** Applicability of landscape metrics for the monitoring of landscape change: issues of scale, resolution and interpretability. *Ecological Indicators* 2, 3-15.
- Lechner, A., M., Langford, W., T., Bekessy, S., A., Jones, S., D., 2012.** Are landscape ecologists addressing uncertainty in their remote sensing data? *Landscape Ecology* 27(9), 1249-1261.
- Lechner, A., M., Reinke, K., J., Wang, Y., Bastin, L., 2013.** Interactions between landcover pattern and geospatial processing methods: Effects on landscape metrics and classification accuracy. *Ecological Complexity* 15, 71-82.
- Leitão, A., B., Miller, J., Ahern, J., McGarigal, K., 2006.** Measuring landscapes: a planner's handbook. Island Press, Washington DC, 245.
- Li, H., Wu., 2004.** Use and misuse of landscape indices. *Landscape Ecology* 19(4), 389-399.
- Lu, D., Weng, Q., 2007.** A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing* 28(5), 823-870.
- Mas, J., F., Gao, Y., Navarrete Pacheco, J., A., 2010.** Sensitivity of landscape pattern metrics to classification approaches. *Forest Ecology and Management* 259(7), 1215-1224.
- McGarigal, K., Marks, B., 1995.** FRAGSTATS: Spatial Analysis Program for Quantifying Landscape Structure. USDA Forest Service General Technical Report PNW-GTR-351. USDA Forest Service, Washington DC.

McGarigal, K., 2002. Landscape pattern metrics. In: A.H. El-Shaarawi and W. W. Piegorsch, (Editors), Encyclopedia of Environmetrics. John Wiley & Sons, Chichester, England, 1135-1142.

McGarigal, K., SA Cushman, and E Ene. 2012. FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Saadaval internetiaadressil: <http://www.umass.edu/landeco/research/fragstats/fragstats.html>

Peña, J., M., Lozano, J., A., Larranaga, P., 1999. An empirical comparison of four initialization methods for the k-means algorithm. Pattern Recognition Letters 20(10), 1027–1040.

Peng, J., Wang, Y., Ye, M., Wu, J., Zhang, Y., 2007. Effects of land-use categorization on landscape metrics: A case study in urban landscape of Shenzhen, China. International Journal of Remote Sensing 28(21), 4877-4895.

RapidEye Standard Image Product Specifications 2013. BlackBridge.

Redmond, S., J., Heneghan, C., 2007. A method for initialising the K-means clustering algorithm using *kd*-trees. Pattern Recognition Letters 28(8), 965–973.

Riitters, K., H., O'Neill, R., V., Hunsaker, C., T., Wickham, J., D., Yankee, D., H., Timmins, S., P., Jones, K., B., Jackson, B., L., 1995. A factor-analysis of landscape pattern and structure metrics. Landscape Ecology 10(1), 23-39.

Saura, S., Martinez-Millan, J., 2001. Sensitivity of landscape pattern metrics to map spatial extent. Photogrammetric Engineering & Remote Sensing 67, 1027–1036.

Saura, S., 2004. Effects of remote sensor spatial resolution and data aggregation on selected fragmentation indices. Landscape Ecology 19(2), 197-209.

Shao, G., Liu, D., Zhao, G., 2001. Relationships of image classification accuracy and variation of landscape statistics. Canadian Journal of Remote Sensing 27(1), 35-45.

Shao, G., Wu, J., 2008. On the accuracy of landscape pattern analysis using remote sensing data. Landscape Ecology 23(5), 505-511.

Símová, P., Gdulová, K., 2012. Landscape indices behavior: A review of scale effects. Applied Geography 34, 385-394.

Steele, B.M., Winne, J.C., Redmond, R.L., 1998. Estimation and mapping of misclassification probabilities for thematic land cover maps. Remote Sensing of Environment 66, 192–202.

Turner, M., G., 1989. Landscape ecology – the effect of pattern on process. Annual Review of Ecology and Systematics 20, 171-197.

Uuemaa, E., Roosaare, J., Mander, Ü., 2008. Maastikuindeksid biogeensete ainete väljakande indikaatorina – kaugseire ja ruumiandmete kasutusvõimalused. Rmt. Väljataga, K., Kaukver, K. (toim), Kaugseire Eestis. Keskkonnainfo, Tallinn, 132-143.

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., Mander, Ü., 2009. Landscape Metrics and Indices: An Overview of Their Use in Landscape Research. *Living Reviews in Landscape Research* 3(1).

Uuemaa, E., Roosaare, J., Oja, T., Mander, Ü., 2011. Analysing the structure of the Estonian landscapes: which landscape metrics are the most suitable for comparing different landscapes? *Estonian Journal of Ecology* 60(1), 70-80.

Uuemaa, E., Mander, Ü., Marja, R., 2012. Trends in the use of landscape spatial metrics as landscape indicators: A review. *Ecological Indicators* 28, 100–106.

Wu, J., 2004. Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology* 19(2), 125-138.

Lisad

Lisa 1. Töös kasutatavad programmi Fragstats maastikku kirjeldavad indeksid (McGarigal *et al.* 2012)

Programmis Fragstats on kahte liiki klassi ja maastiku taseme indekseid: esiteks kompositsiooni ja konfiguratsiooni indeksid, teiseks indeksite statistilist jaotust kirjeldavad indeksid (keskmine **MN** - *mean*, pindalaga kaalutud keskmine **AM** - *area-weighted mean*, mediaan **MD** - *median*, ulatus ehk suurima ja väikseima vahe **RA** - *range*, standardhälve **SD** - *standard deviation*, variatsioonikoefitsient **CV** - *coefficient of variation*). Keskmise ja pindalaga kaalutud keskmise põhiline erinevus on see, et esimene on eraldistekeskne, teine maastikukeskne, iseloomustades juhuslikult valitud ala maastikul või pikslit satelliidipildil.

Pindala ja serva indeksid (*Area & edge metrics*):

LPI (*Largest Patch Index*) – suurima eraldise indeks, näitab suurima eraldise protsentuaalset osakaalu maastikus.

$$LPI = \frac{\max(a_{ij})}{A} (100)$$

kus a_{ij} = eraldise ij pindala (m^2);

A = maastiku kogupindala (m^2).

TE (*total edge*) – kogu serva pikkus. Servaks loetakse piiri kahe eri maastikuklassi eraldiste vahel.

$$TE = E$$

kus E = kogu serva pikkus (m) maastikus.

ED (*Edge Density*) – servatihedus, kõikide eraldiste serva pikkuste summa jagatud maastiku pindalaga. Arvutatakse meetrites hektari kohta.

$$ED = \frac{E}{A} (10\,000)$$

kus E = kogu serva pikkus (m) maastikus;

A = kogu maastiku pindala (m^2).

AREA_MN, AREA_AM, AREA_MD, AREA_RA, AREA_SD, AREA_CV (*Patch Area Distribution*) – eraldiste pindalalist jaotumist kirjeldavad indeksid: keskmine pindala, pindalaga kaalutud keskmine pindala, mediaan pindala, pindalaline ulatus (suurima ja väikseima pindala vahe), pindalade standardhälve, pindalade variatsioonikoefitsient. Arvutatakse hektarites.

$$AREA = a_{ij} \left(\frac{1}{10\,000} \right)$$

kus a_{ij} = eraldise ij pindala (m²).

GYRATE_MN, GYRATE_AM, GYRATE_MD, GYRATE_RA, GYRATE_SD, GYRATE_CV (*Radius of Gyration Distribution*) – pöörlemise raadiust ehk eraldiste ulatust maastikus mõõtvad indeksid. GYRATE väärtus sõltub vahemaast eraldise kõikide pikslite/punktide ja keskpunkti vahel; mida suuremad ja vähem kompaktsamad eraldised, seda suurem väärtus. Arvutatakse meetrites.

$$GYRATE = \sum_{r=1}^z \frac{h_{ijr}}{z}$$

kus h_{ijr} = kaugus (m) eraldise punkti/piksli ijr ja eraldise ij keskpunkti vahel;
z = punktide / pikslite arv eraldise ij sees.

Kuju indeksid (*Shape metrics*):

PAFRAC (*Perimeter-Area Fractal Dimension*) – ümbermõõdu-pindala fraktaalne dimensioon, mõõdab eraldiste kuju keerukust eri skaalades (erineva suurusega eraldiste lõikes). Väärtused on vahemikus 1 kuni 2, kus väärtus 1 viitab ruudukujulisusele (lihtne kuju) ja väärtus 2 keerukale käänulisele kujule.

$$PAFRAC = \frac{2}{\frac{[N \sum_{i=1}^m \sum_{j=1}^n (\ln p_{ij} \cdot \ln a_{ij})] - [(\sum_{i=1}^m \sum_{j=1}^n \ln p_{ij})(\sum_{i=1}^m \sum_{j=1}^n \ln a_{ij})]}{(N \sum_{i=1}^m \sum_{j=1}^n \ln p_{ij}^2) - (\sum_{i=1}^m \sum_{j=1}^n \ln p_{ij})^2}}$$

kus p_{ij} = eraldise ij ümbermõõt;

a_{ij} = eraldise ij pindala (m²);

N= eraldiste koguarv maastikus.

PARA_MN, PARA_AM, PARA_MD, PARA_RA, PARA_SD, PARA_CV (*Perimeter-Area Ratio Distribution*) – ümbermõõdu-pindala suhet kirjeldavad indeksid. Sõltuvad eraldiste suurusest, mistõttu ei ole kõige paremad kuju kirjeldajad.

$$PARA = \frac{p_{ij}}{a_{ij}}$$

kus p_{ij} = eraldise ij ümbermõõt;

a_{ij} = eraldise ij pindala (m²).

SHAPE_MN, SHAPE_AM, SHAPE_MD, SHAPE_RA, SHAPE_SD, SHAPE_CV (*Shape Index Distribution*) – kujuindeksid, mõõdavad eraldiste kuju keerukust, võrreldes sama pindalaga ruuduga, mistõttu leevendavad indeksi PARA probleemi – sõltuvust pindalast. Väärtused algavast 1-st, ülempiir puudub.

$$\text{SHAPE} = \frac{.25 p_{ij}}{\sqrt{a_{ij}}}$$

kus p_{ij} = eraldise ij ümbermõõt (m);
 a_{ij} = eraldise ij pindala (m²).

FRAC_MN, FRAC_AM, FRAC_MD, FRAC_RA, FRAC_SD, FRAC_CV (*Fractal Index Distribution*) – fraktaalset dimensiooni ehk ümbermõõdu-pindala logaritmilist suhet kirjeldavad indeksid. Nagu kujuindeks SHAPE, kirjeldavad kuju keerukust, sõltumata nii palju eraldise pindalast kui PARA. Väärtused on vahemikus 1 kuni 2, kus suurem väärtus viitab keerukamale kujule.

$$\text{FRAC} = \frac{2 \ln(.25 p_{ij})}{\ln a_{ij}}$$

kus p_{ij} = eraldise ij ümbermõõt (m);
 a_{ij} = eraldise ij pindala (m²).

CIRCLE_MN, CIRCLE_AM, CIRCLE_MD, CIRCLE_RA, CIRCLE_SD, CIRCLE_CV (*Related Circumscribing Circle Distribution*) – väikseima võimaliku ümbritseva ringi indeksid, sobivad leidmaks eraldisi, mis on korruga piklikud ja välja venitatud. Väärtused on vahemikus 0 kuni 1, kus 1 viitab piklikule, lineaarsele eraldise kujule.

$$\text{SQUARE} = 1 - \left[\frac{a_{ij}}{a_{ij}^s} \right]$$

kus a_{ij} = eraldise ij ümbermõõt (m);
 a_{ij}^s = väikseima eraldist ij ümbritseva ringi pindala (m²).

CONTIG_MN, CONTIG_AM, CONTIG_MD, CONTIG_RA, CONTIG_SD, CONTIG_CV (*Contiguity Index Distribution*) – pidevust kirjeldavad indeksid, mõõdavad kuju kompaktsust ehk kui hästi on eraldise pikslid omavahel ühendatud. Väärtused arvutatakse vahemikus 0 kuni 1.

$$\text{CONTIG} = \frac{\left[\frac{\sum_{r=1}^Z c_{ijr}}{a_{ij}^*} \right] - 1}{v - 1}$$

kus c_{ijr} = eraldise ij piksli r pidevuse väärtus, $r = 1$, kui piksel kuulub eraldisele, ja $r = 0$, kui ei kuulu;

v = pidevuse väärtuste summa 3x3 pikslite liikuva akna järgi, ortogonaalselt paiknevad eraldise ij pikslid saavad väärtuse 2, diagonaalselt paiknevad 1, eraldisele ij mitte kuuluvad pikslid väärtustatakse 0-ga;

a_{ij}^* = eraldise ij pindala pikslite arvu järgi.

Tuumala indeksid (*Core area metrics*):

TCA (*Total Core Area*) - tuumala pindala hektarites, eelnevalt peab olema defineeritud servasügavus, mis määrab, kui suure ala eraldisest moodustab serv.

$$\text{TCA} = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^c \left(\frac{1}{10\,000} \right)$$

kus a_{ij}^c = eraldise ij tuumala (m^2), sõltuvalt servasügavusest (m).

NDCA (*Number of Disjunct Core Areas*) – liigendatud tuumalade arv maastikus – ühe eraldise sees võib olla mitu tuumala, kui eraldis on näiteks liigendatud kujuga. NDCA = 0, kui kõik alad eraldiste sees jäävad servasügavusega määratud piirkondadesse.

$$\text{NDCA} = \sum_{i=1}^m \sum_{j=1}^n n_{ij}^c$$

kus n_{ij}^c = eraldise ij tuumalade arv, sõltuvat servasügavusest (m)

DCAD (*Disjunct Core Area Density*) – liigendatud tuumalade tihedus, liigendatud tuumalade arv jagatud kogu ala pindalaga. Väärtuseks on tuumalade arv 100 ha kohta.

$$\text{DCAD} = \frac{\sum_{i=1}^m \sum_{j=1}^n n_{ij}^c}{A} (10\,000) (100)$$

kus n_{ij}^c = eraldise ij tuumalade arv, sõltuvat servasügavusest (m);

A = kogu maastiku pindala (m^2)

CORE_MN, CORE_AM, CORE_MD, CORE_RA, CORE_SD, CORE_CV (*Core Area Distribution*) – tuumalade jaotumist kirjeldavad indeksid. Tuumala on ala eraldise sees, mis jääb välja servaefekti mõjualast. Arvutatakse hektarites.

$$\text{CORE} = a_{ij}^c \left(\frac{1}{10\,000} \right)$$

kus a_{ij}^c = eraldise ij tuumala (m²), sõltuvalt servasügavusest (m).

DCORE_MN, DCORE_AM, DCORE_MD, DCORE_RA, DCORE_SD, DCORE_CV (*Disjunct Core Area Distribution*) - liigendatud tuumalade jaotumist kirjeldavad indeksid. Ühe eraldise sees võib olla mitu tuumala, mida arvestatakse eraldi, erinevalt CORE_ indeksitest, mis arvestavad eraldise kõiki tuumalasid tervikuna.

CAI_MN, CAI_AM, CAI_MD, CAI_RA, CAI_SD, CAI_CV – (*Core Area Index Distribution*) – tuumalade osakaal eraldistes. Need indeksid näitavad, kui suure osa protsentuaalselt moodustavad eraldistest tuumalad.

$$\text{CAI} = \frac{a_{ij}^c}{a_{ij}} (100)$$

kus a_{ij}^c = eraldise ij tuumala (m²), sõltuvalt servasügavusest (m);
 a_{ij} = eraldise ij pindala.

Kontrastsuse indeksid (*Contrast metrics*):

CWED (*Contrast-Weighted Edge Density*) – kontrastsusega kaalutud servade tihedus, sõltub servapikkusest ja eri maastikuklassidele antud serva kontrastsuse kaaludest. Arvutatakse meetrites hektari kohta.

$$\text{CWED} = \frac{\sum_{i=1}^m \sum_{k=i+1}^m (e_{ik} \cdot d_{ik})}{A} (10\,000)$$

kus e_{ik} = maastiku kogu serva pikkus (m) eraldiste tüüpide (klasside) i ja k vahel; sisaldab ka eraldise tüübi i maastiku piiri lõike;

d_{ik} = erinevus (kasutaja määratud serva kontrastsuse kaal) eraldiste tüüpide i ja k vahel;

A = kogu maastiku pindala (m²).

TECI (*Total Edge Contrast Index*) – serva kontrastsuse indeks, näitab servade kontrastust kogu maastikus protsentides. Indeksi väärtus on 0, kui maastik koosneb ainult ühest eraldisest (kontrastsus puudub) ja väärtus 100, kui kõikide servade kontrastsus on maksimaalne. Eelnevalt peab olema defineeritud eritüüpiliste maastikuklasside serva kontrastsuse maatriks.

$$TECI = \frac{\sum_{i=1}^m \sum_{k=i+1}^m (e_{ik} \cdot d_{ik})}{E^*} (100)$$

kus e_{ik} = maastiku kogu serva pikkus (m) eraldiste tüüpide (klasside) i ja k vahel; sisaldab ka eraldise tüübi i maastiku piiri lõike;

E^* = kogu maastiku servapikkus (m), sisaldab maastiku piiri;

d_{ik} = erinevus (kasutaja määratud serva kontrastsuse kaal) eraldiste tüüpide i ja k vahel.

ECON_MN, ECON_AM, ECON_MD, ECON_RA, ECON_SD, ECON_CV (*Edge Contrast Index Distribution*) – serva kontrastsuse jagunemist maastikus kirjeldavad indeksid, mõõdavad eraldise ümbermõõdu suhtelist serva kontrastsust. Arvutatakse protsentides.

$$ECON = \frac{\sum_{k=1}^m (p_{ijk} \cdot d_{ik})}{p_{ij}} (100)$$

kus p_{ijk} = eraldise ij servapikkus, mis külgneb eraldise tüübiga (klassiga) k;

d_{ik} = erinevus (kasutaja määratud serva kontrastsuse kaal) eraldiste tüüpide i ja k vahel.

p_{ij} = eraldise ij ümbermõõtu (m).

Koonduvuse indeksid (*Aggregation metrics*):

CONTAG (*Contagion*) – koonduvus, mõõdab ühte tüüpi pikslite naabrussuhete arvu, mõõtes sellega kaudselt eraldiste koondumist / hajumist maastikus. Indeksi väärtus on 0, kui eraldiste tüübid on maksimaalselt hajunud (iga piksel esindab erinevat eraldise tüüpi) ning eraldised vahelduvad maksimaalselt. Indeksi väärtus on suur suurte kompaktsete eraldiste korral. Seotud pöördvõrdeliselt servatihedusega. Arvutatakse protsentides.

$$CONTAG = \left[1 + \frac{\sum_{i=1}^m \sum_{k=1}^m \left[P_i \cdot \frac{g_{ik}}{\sum_{k=1}^m g_{ik}} \right] \cdot \ln \left[P_i \cdot \frac{g_{ik}}{\sum_{k=1}^m g_{ik}} \right]}{2 \ln(m)} \right] (100)$$

kus P_i = eraldise tüübi (klassi) i poolt hõivatud maastiku osakaal;

g_{ik} = eraldiste tüüpide (klasside) i ja k pikslite vahelised külgnevused

m = eraldiste tüüpide (klasside) arv maastikus.

IJI (*Interspersion & Juxtaposition Index*): vahelduvuse ja kõrvuti paiknemise indeks, põhineb eraldiste (mitte pikslite, nagu koonduvuse indeks CONTAG) külgnevuste suhetel. Indeksi väärtus on maksimaalne, kui kõik maastikutüübid vahelduvad üksteisega võrdselt. Arvutatakse protsentides.

$$IJI = \frac{-\sum_{i=1}^m \sum_{k=i+1}^m \left[\left(\frac{e_{ik}}{E} \right) \cdot \ln \left(\frac{e_{ik}}{E} \right) \right]}{\ln(0.5[m(m-1)])} \quad (100)$$

kus e_{ik} = eraldiste tüüpide (klasside) i ja k vahelise kogu serva pikkus (m) maastikus;
 E = kogu serva pikkus (m) maastikus, v.a. tagataust;
 m = eraldiste tüüpide (klasside) arv maastikus.

PLADJ (*Percentage of Like Adjacencies*) – sama tüüpi naabrussuhete protsent. Võrdub 0, kui sama tüüpi eraldised on maksimaalselt hajutatud, ja 100, kui maksimaalselt agregeerunud (maastik koosneb ühest eraldisest). Suuremate ja lihtsa kujuga eraldistega maastikus on suurem sama tüüpi naabrussuhete osakaal. Erinevalt koonduvusest, mõõdab see indeks ainult hajuvust, mitte vahelduvust.

$$PLADJ = \left(\frac{\sum_{i=1}^m (g_{ii})}{\sum_{i=1}^m \sum_{k=1}^m (g_{ijk})} \right) \quad (100)$$

kus g_{ii} = eraldise tüübi (klassi) i tema enda tüüpi kõrvuti asetsevate pikslite arv;
 g_{ik} = eraldiste tüüpide (klasside) i ja k pikslite vahelised külgnevused.

AI (*Aggregation Index*) – agregatsiooni indeks, näitab sama maakattetüübi pikslite koonduvust maastikus. Arvutatakse protsentides.

$$AI = \left[\sum_{i=1}^m \left(\frac{g_{ii}}{\max \rightarrow g_{ii}} \right) P_i \right] \quad (100)$$

kus g_{ii} = eraldise tüübi (klassi) i tema enda tüüpi kõrvuti asetsevate pikslite arv;
 $\max \rightarrow g_{ii}$ = eraldise tüübi (klassi) i tema enda tüüpi kõrvuti asetsevate pikslite maksimaalne arv;
 P_i = eraldise tüübi (klassi) i poolt hõivatud maastiku osakaal.

LSI (*Landscape Shape Index*) – maastiku kuju indeks, eraldiste kogu serva pikkust võrreldakse ristkülikuga, jagades kogu serva pikkuse vähima võimaliku servapikkusega.

$$LSI = \frac{.25 E^*}{\sqrt{A}}$$

kus E^* = kõikide servade kogupikkus (m) maastikus;
 A = maastiku kogupindala (m²).

COHESION (*Patch Cohesion Index*) – eraldise sidususe indeks, mis klassi tasemel mõõdab vastava eraldise sidusust, võttes arvesse eraldise ümbermõõtu ja pindalat, kuid maastiku tasemel ei ole selle indeksi käitumist veel hinnatud. Indeksi väärtuste ulatus ei ole teada.

$$COHESION = \left[1 - \frac{\sum_{i=1}^m \sum_{j=1}^n p_{ij}^*}{\sum_{i=1}^m \sum_{j=1}^n p_{ij}^* \sqrt{a_{ij}^*}} \right] \cdot \left[1 - \frac{1}{\sqrt{Z}} \right]^{-1} \cdot (100)$$

kus p_{ij}^* = eraldise ij ümbermõõt eraldise pikslite arvu järgi;

a_{ij}^* = eraldise ij pindala eraldise pikslite arvu järgi;

Z = pikslite koguarv maastikus.

NP (*Number of Patches*) – eraldiste arv maastikus

$$NP = N$$

kus N = eraldiste koguarv maastikus.

PD (*Patch Density*) – eraldiste tihedus, eraldiste arv jagatud maastiku pindalaga. Eraldiste arv arvutatakse 100 ha kohta.

$$PD = \frac{N}{A} (10\,000)(100)$$

kus N = eraldiste koguarv maastikus;

A = maastiku kogupindala (m²).

DIVISION (*Landscape Division Index*) – maastiku eraldisteks jagunemise indeks, põhineb tõenäosusel, et kaks juhuslikult valitud pikslit maastikul ei kuulu samale eraldisele. Arvutatakse vahemikus 0 kuni 1, kus maksimaalne väärtus viitab sellele, et iga piksel kuulub erinevale eraldisele ehk moodustab iseseisva eraldise.

$$DIVISION = 1 - \left[\sum_{i=1}^m \sum_{j=1}^n \left(\frac{a_{ij}}{A} \right)^2 \right]$$

kus a_{ij} = eraldise ij pindala (m²);

A = maastiku kogupindala (m²).

SPLIT (*Splitting Index*) – tükeldatuse indeks, väljendab maastiku tükeldatust, mis sõltub eraldiste arvust ja pindalast; väärtus kasvab maastiku jagunemisel väikeste eraldiste vahel, maksimaalne väärtus saavutatakse siis, kui iga piksel moodustab omaette eraldise.

$$\text{SPLIT} = \frac{A^2}{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

kus a_{ij} = eraldise ij pindala (m^2);

A = maastiku kogupindala (m^2).

MESH (*Effective Mesh Size*) – koondatuse indeks, põhineb eraldiste suurusel. Maksimaalne väärtus saavutatakse siis, kui maastik koosneb vaid ühest eraldisest ja minimaalne, kui iga piksel moodustab omaette eraldise. Arvutatakse hektarites.

$$\text{MESH} = \frac{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}{A}$$

kus a_{ij} = eraldise ij pindala (m^2);

A = maastiku kogupindala (m^2).

ENN_MN, ENN_AM, ENN_MD, ENN_RA, ENN_SD, ENN_CV (*Euclidean Nearest Neighbor Distance Distribution*) – eukleidiline lähima naabri kauguse indeksite rühm, kus arvutatakse lühimad kaugused sama tüüpi maastikuklassi eraldiste vahel. Need indeksid mõõdavad eraldiste isoleeritust. Väärtused antakse meetrites.

$$\text{ENN} = h_{ij}$$

kus h_{ij} = eraldise ij kaugus (m) lähima sama maakatte/-kasutuse tüübiga naabereraldisest.

PROX_MN, PROX_AM, PROX_MD, PROX_RA, PROX_SD, PROX_CV (*Proximity Index Distribution*) – läheduse indeksite rühm, näitab sama tüüpi eraldiste arvukust kasutaja poolt määratud otsinguraadiuse sees. Võtab arvesse nii eraldiste suurust kui ka lähedust otsimisraadiuses. Indeksitel ei ole ühikuid, mistõttu saab neid kasutada vaid väärtuste võrdlemiseks.

$$\text{PROX} = \sum_{s=1}^n \frac{a_{ijs}}{h_{ijs}^2}$$

kus a_{ijs} = eraldise ijs pindala (m^2) eraldise ij määratletud naabruskonnas (m);

h_{ijs} = eraldiste ijs ja ijs (sama tüüpi eraldiste) servade vaheline kaugus.

SIMI_MN, SIMI_AM, SIMI_MD, SIMI_RA, SIMI_SD, SIMI_CV (*Similarity Index Distribution*) – sarnasuse indeksite rühm. Sarnasuse indeks on läheduse indeksi

modifikatsioon, võtab arvesse eraldiste suurust, lähedust ja sarnasust määratud otsinguraadiuse sees, sõltumata maakatte tüübist. Eelnevalt peab kasutaja määrama eri tüüpi maastikuklassidele sarnasuse kaalud ehk sarnasuse maatrksi. Indeksil ei ole ühikuid.

$$\text{SIMI} = \sum_{s=1}^n \frac{a_{ijs} d_{ik}}{h_{ijs}^2}$$

kus a_{ijs} = eraldise ijs pindala (m^2) eraldise ij määratletud naabruskonnas (m);
 d_{ik} = sarnasus (kasutaja määratud sarnasuse kaal) eraldiste tüüpide i ja k vahel;
 h_{ijs} = eraldiste ijs ja ijs (sama tüüpi eraldiste) servade vaheline kaugus.

CONNECT (*Connectance Index*) – sidususe indeks, näitab protsenti võimalikust maksimaalsest sidususest antud eraldiste arvu juures. Põhineb kasutaja poolt määratud kaugusel. Väärtused antakse protsentides.

$$\text{CONNECT} = \left[\frac{\sum_{i=1}^m \sum_{j \neq k}^m c_{ijk}}{\sum_{i=1}^m \frac{n_i(n_i - 1)}{2}} \right] (100)$$

kus c_{ijk} = sama maastikutüübi (klassi) eraldiste j ja k vaheline ühendatus (0 = ühendus puudub, 1 = omavahel ühenduses), sõltub kasutaja poolt määratud lävendkaugusest;
 n_i = iga maastikutüübi (klassi) i eraldiste arv.

Mitmekesisuse indeksid (*Diversity metrics*):

SHDI (*Shannon's Diversity Index*) – Shannoni mitmekesisuse indeks, hindab maastiku keerukust, mis arvutatakse selle määramatuste summana. Indeks on tundlik eraldiste rohkuse ja harva esinevate eraldiste tüüpide suhtes. Mida rohkem on maastikus eri tüüpi eraldisi ja mida võrdsema osa nad maastikust hõlmavad, seda suurem on indeksi väärtus. Väärtused algavad 0-st, ülempiir puudub.

$$\text{SHDI} = - \sum_{i=1}^m P_i \cdot \ln P_i$$

kus P_i = eraldise tüübi (klassi) i osakaal maastikus.

SIDI (*Simpson's Diversity Index*) – Simpsoni mitmekesisuse indeks, mille väärtus näitab tõenäosust, et kaks juhuslikult valitud pikslit kuuluvad eri tüüpi eraldistele. Võrreldes Shannoni mitmekesisuse indeksiga, on indeks vähem tundlik harva esinevate eraldiste tüüpide suhtes. Väärtused jäävad vahemikku 0 kuni 1.

$$\text{SIDI} = 1 - \sum_{i=1}^m P_i^2$$

kus P_i = eraldise tüübi (klassi) i osakaal maastikus.

MSIDI (*Modified Simpson's Diversity Index*) – muudetud Simpsoni mitmekesisuse indeks. Võrreldes SIDI-ga, sarnaneb indeks rohkem Shannoni mitmekesisuse indeksiga: väärtused algavad 0-st, ülempiiri ei ole.

$$MSIDI = -\ln \sum_{i=1}^m P_i^2$$

kus P_i = eraldise tüübi (klassi) i osakaal maastikus.

SHEI (*Shannon's Evenness Index*) – Shannoni ühtluse indeks, hindab eraldiste tüüpide pindalade jagunemise ühtlust. Väärtused jäävad vahemikku 0 kuni 1, kus 1 viitab maksimaalsele eraldiste tüüpide pindalalisele ühtlusele.

$$SHEI = \frac{-\sum_{i=1}^m (P_i \cdot \ln P_i)}{\ln m}$$

kus P_i = eraldise tüübi (klassi) i osakaal maastikus;
m = eraldise tüüpide (klasside) arv maastikus.

SIEI (*Simpson's Evenness Index*) – Simpsoni ühtluse indeks, Väärtused jäävad vahemikku 0 kuni 1, kus 1 viitab maksimaalsele eraldiste tüüpide pindalalise jagunemise ühtlusele.

$$SIEI = \frac{1 - \sum_{i=1}^m P_i^2}{1 - \left(\frac{1}{m}\right)}$$

kus P_i = eraldise tüübi (klassi) i osakaal maastikus;
m = eraldise tüüpide (klasside) arv maastikus.

MSIEI (*Modified Simpson's Evenness Index*) – muudetud Simpsoni ühtluse indeks, Väärtused jäävad vahemikku 0 kuni 1, kus 1 viitab maksimaalsele eraldiste tüüpide pindalalise jagunemise ühtlusele.

$$MSIEI = \frac{-\ln \sum_{i=1}^m P_i^2}{\ln m}$$

kus P_i = eraldise tüübi (klassi) i osakaal maastikus;
m = eraldise tüüpide (klasside) arv maastikus

Lisa 2. Programmis Fragstats kasutatud eri maastikutüüpide sarnasust ja serva kontrastsust kirjeldavad maatriksid

Kose sarnasuse maatriks etalonidega ja etalonideta klassifitseerimise jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5, 6)

1, 0.6, 0.6, 0.2, 0.2, 0.2
0.6, 1, 0.6, 0.2, 0.2, 0.2
0.6, 0.6, 1, 0.2, 0.2, 0.2
0.2, 0.2, 0.2, 1, 0.6, 0.6
0.2, 0.2, 0.2, 0.6, 1, 0.6
0.2, 0.2, 0.2, 0.6, 0.6, 1

Kose serva kontrastsuse maatriks etalonidega ja etalonideta klassifitseerimise jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5, 6)

0, 0.2, 0.2, 0.6, 0.6, 0.6
0.2, 0, 0.2, 0.6, 0.6, 0.6
0.2, 0.2, 0, 0.6, 0.6, 0.6
0.6, 0.6, 0.6, 0, 0.2, 0.2
0.6, 0.6, 0.6, 0.2, 0, 0.2
0.6, 0.6, 0.6, 0.2, 0.2, 0

Tartu sarnasuse maatriks etalonidega ja etalonideta klassifitseerimise jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5)

1, 0.5, 0.5, 0.1, 0.1
0.5, 1, 0.8, 0.5, 0.5
0.5, 0.8, 1, 0.5, 0.5
0.1, 0.5, 0.5, 1, 0.8
0.1, 0.5, 0.5, 0.8, 1

Tartu serva kontrastsuse maatriks etalonidega ja etalonideta klassifitseerimise jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5)

0, 0.5, 0.5, 0.9, 0.9
0.5, 0, 0.2, 0.5, 0.5
0.5, 0.2, 0, 0.5, 0.5
0.9, 0.5, 0.5, 0, 0.2
0.9, 0.5, 0.5, 0.2, 0

Kose sarnasuse maatriks objektorienteeritud klassifitseerimisega saadud maakatteklasside jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5)

1, 0.9, 0.6, 0.6, 0.1
0.9, 1, 0.6, 0.6, 0.1
0.6, 0.6, 1, 0.9, 0.1
0.6, 0.6, 0.9, 1, 0.1
0.1, 0.1, 0.1, 0.1, 1

Kose serva kontrastsuse maatriks objektorienteeritud klassifitseerimisega saadud maakatteklasside jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5)

0, 0.1, 0.4, 0.4, 0.9

0.1, 0, 0.4, 0.4, 0.9

0.4, 0.4, 0, 0.1, 0.9

0.4, 0.4, 0.1, 0, 0.9

0.9, 0.9, 0.9, 0.9, 0

Tartu sarnasuse maatriks objektorienteeritud klassifitseerimisega saadud maakatteklasside jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5)

1, 0.7, 0.4, 0.4, 0.5

0.7, 1, 0.6, 0.6, 0.1

0.4, 0.6, 1, 0.9, 0.1

0.4, 0.6, 0.9, 1, 0.1

0.5, 0.1, 0.1, 0.1, 1

Tartu serva kontrastsuse maatriks objektorienteeritud klassifitseerimisega saadud maaklasside jaoks:

CLASS_LIST_NUMERIC(1, 2, 3, 4, 5)

0, 0.3, 0.6, 0.6, 0.5

0.3, 0, 0.4, 0.4, 0.9

0.6, 0.4, 0, 0.1, 0.9

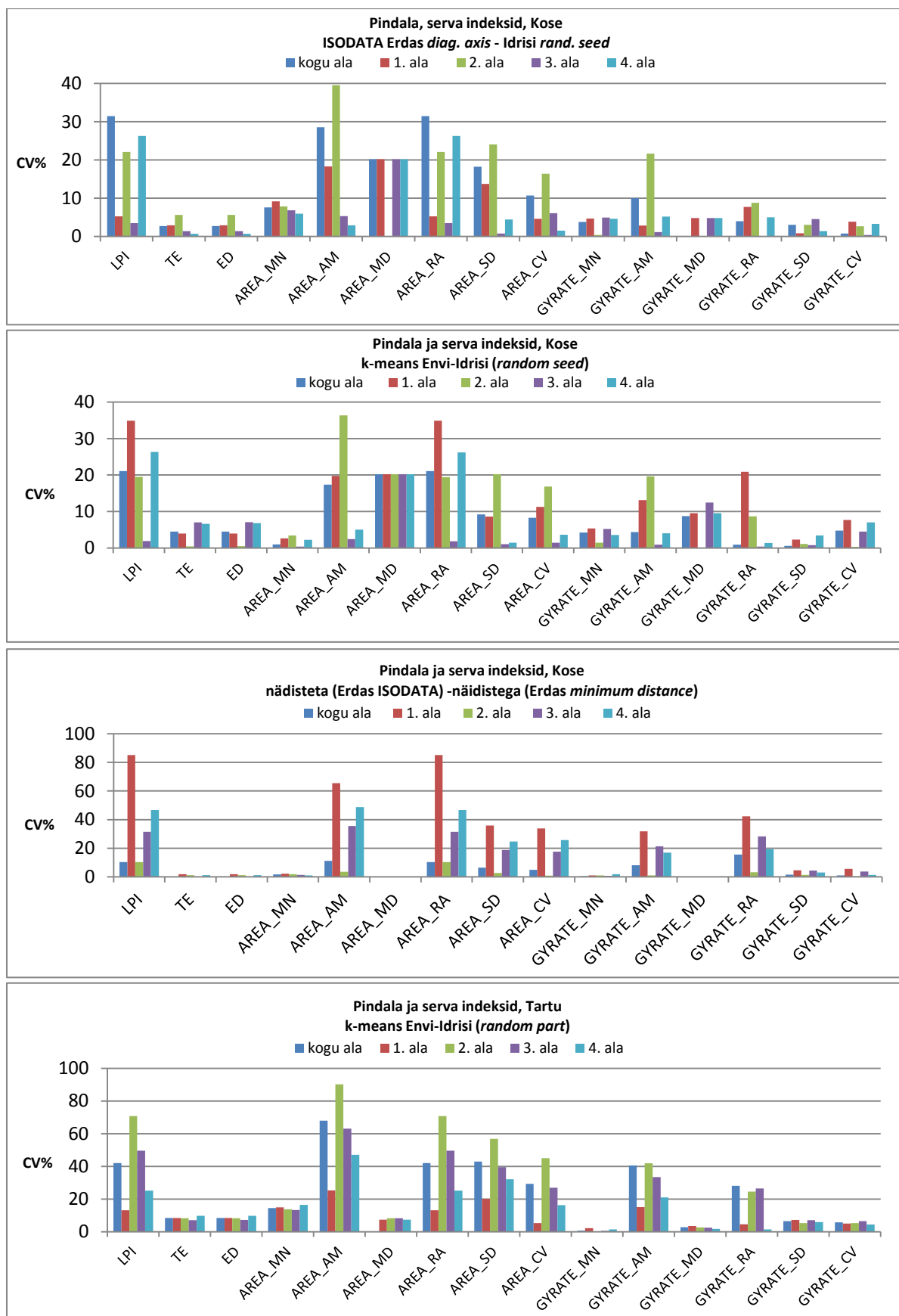
0.6, 0.4, 0.1, 0, 0.9

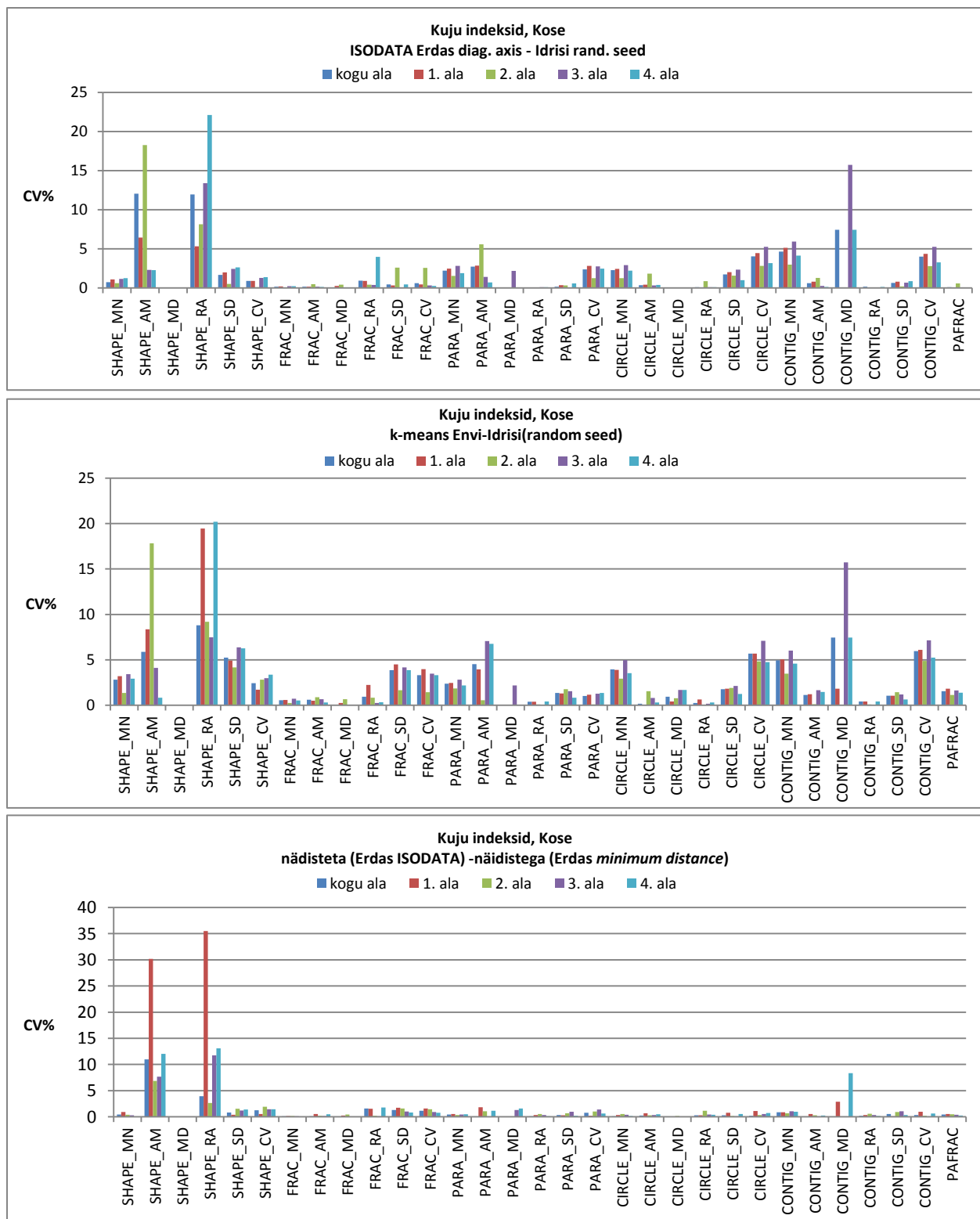
0.5, 0.9, 0.9, 0.9, 0

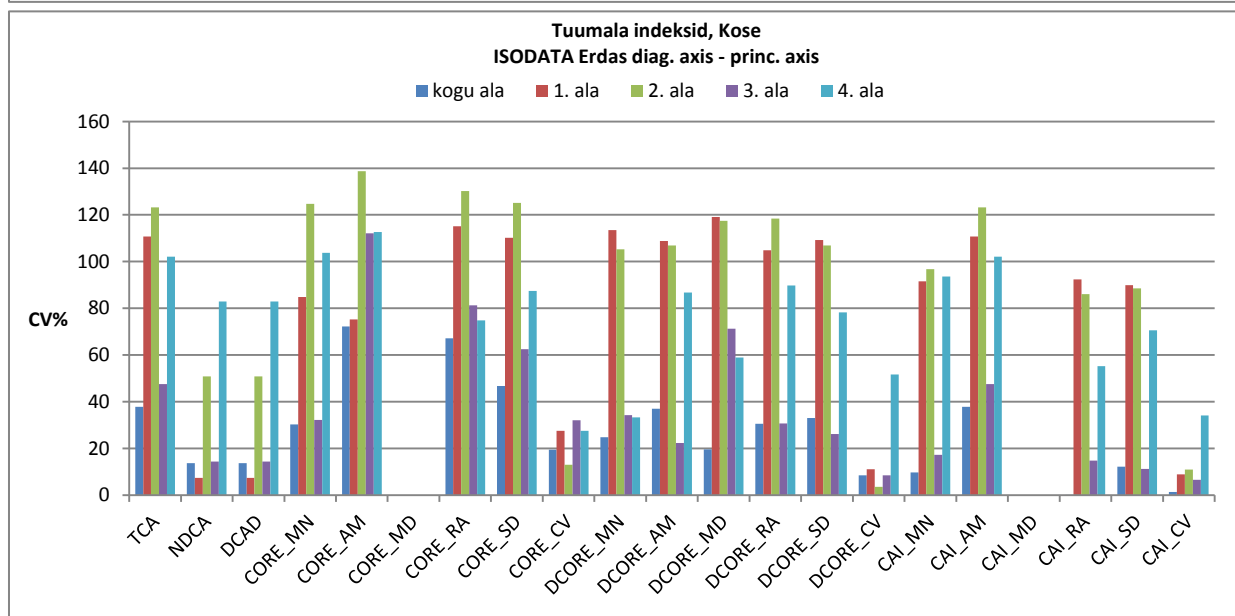
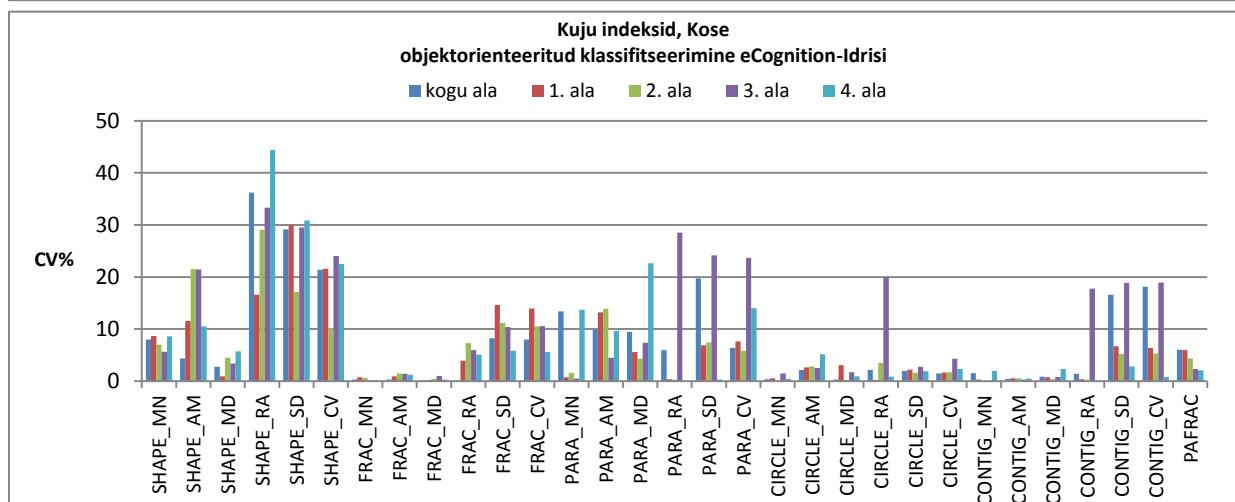
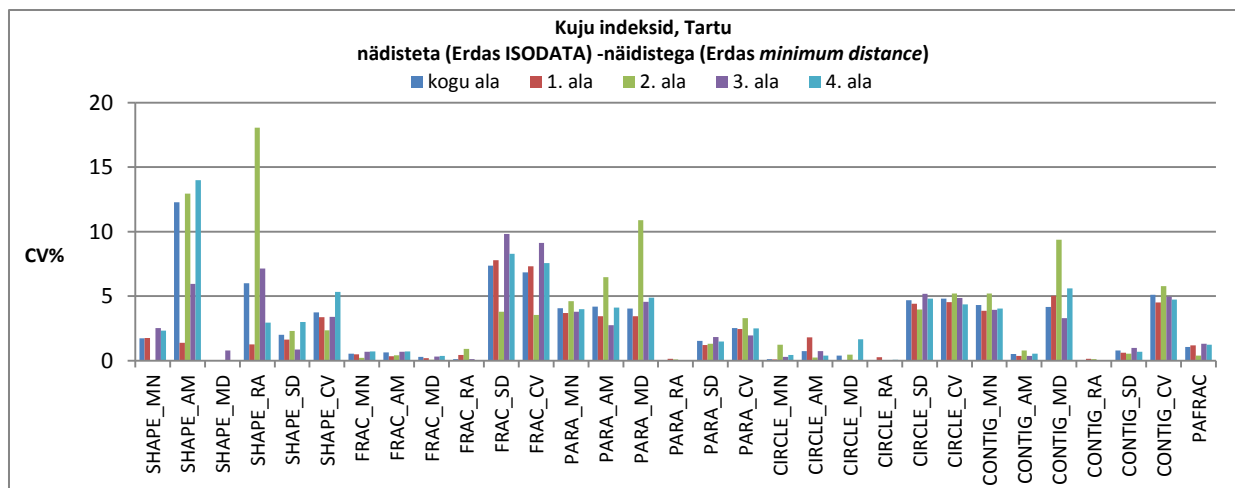
Lisa3. Maastiku taseme indeksite kõikide alade keskmised ja mediaan variatsioonikoefitsendid

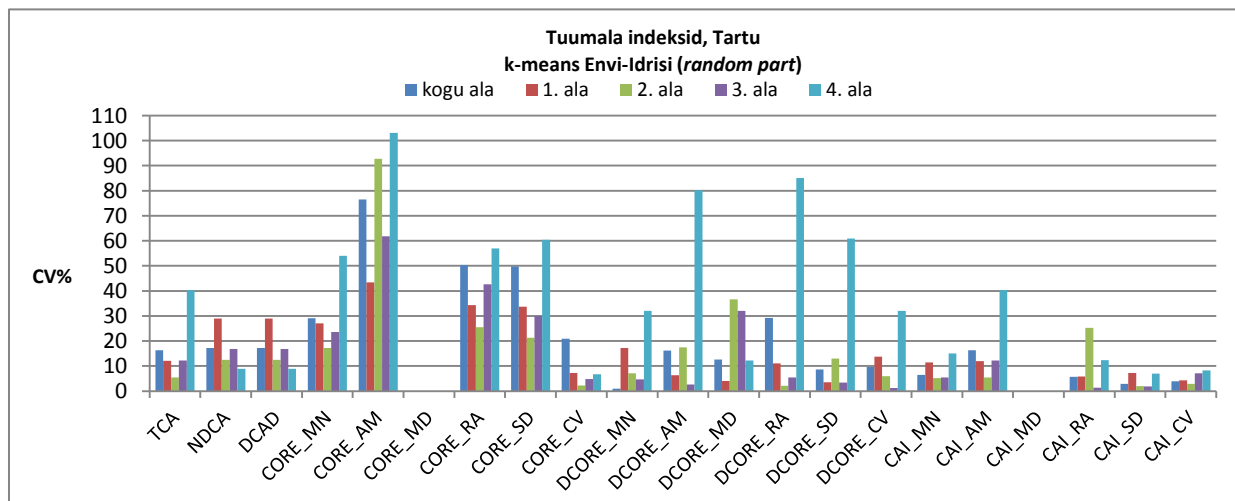
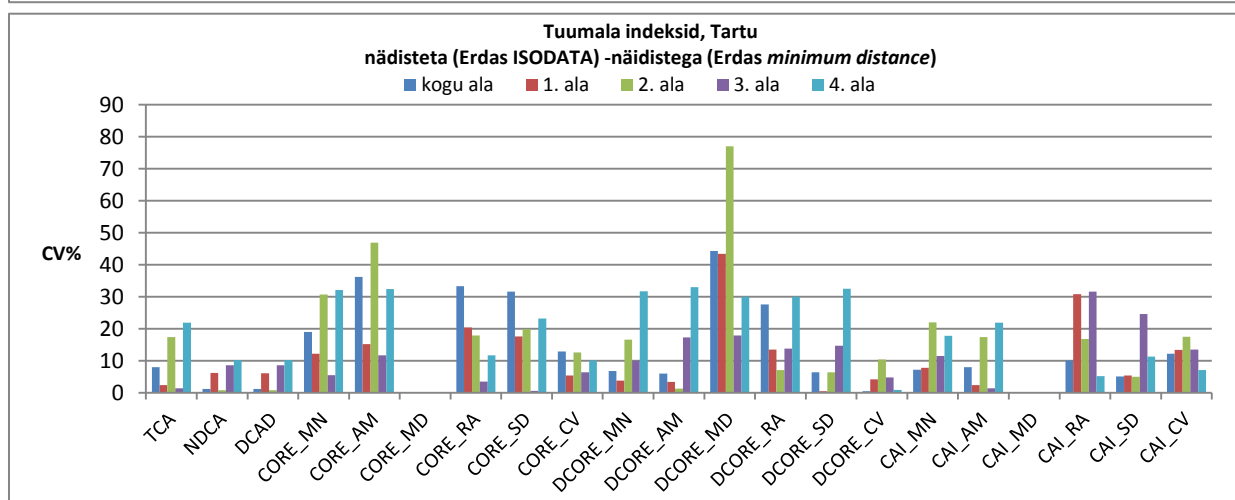
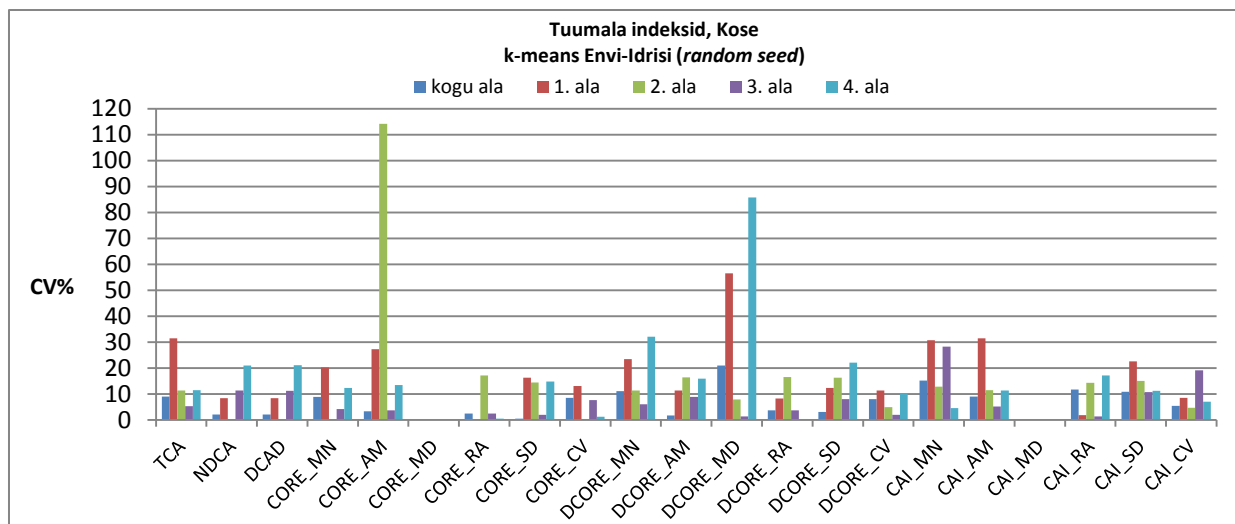
Indeks	CV% keskmine	CV% mediaan	Indeks	CV% keskmine	CV% mediaan	Indeks	CV% keskmine	CV% mediaan
CORE_MD	0.00	0.00	CONTAG	4.63	3.45	CWED	14.17	8.27
CAI_MD	0.00	0.00	ECON_SD	4.67	3.65	GYRATE_RA	14.48	10.44
COHESION	0.16	0.13	ENN_CV	4.75	3.22	AREA_MD	15.70	19.73
FRAC_MD	0.51	0.24	GYRATE_MN	4.75	3.68	SHAPE_AM	15.73	11.72
FRAC_MN	0.58	0.34	MSIDI	4.90	2.49	AREA_CV	16.64	10.97
CIRCLE_MD	0.61	0.40	MSIEI	4.90	2.49	CAI_RA	16.68	10.88
AI	0.80	0.53	TECI	5.49	2.13	GYRATE_AM	17.48	16.00
PLADJ	0.80	0.53	GYRATE_MD	5.70	4.53	SHAPE_RA	18.56	13.30
CONTIG_AM	0.91	0.58	CONTIG_SD	5.70	0.85	CAI_SD	18.63	12.88
FRAC_AM	0.93	0.62	PARA_MD	5.88	2.73	CAI_AM	22.87	12.97
CIRCLE_AM	1.36	0.74	GYRATE_SD	6.12	5.31	TCA	22.88	12.99
ECON_RA	1.36	0.53	ECON_AM	6.22	2.35	DCORE_MN	23.52	16.56
CIRCLE_MN	1.39	0.62	PARA_MN	6.23	2.80	AREA_SD	23.53	19.43
SHAPE_MD	1.60	0.00	FRAC_CV	6.24	4.47	CAI_MN	24.60	17.06
PAFRAC	1.64	1.09	PARA_SD	6.32	1.02	PROX_MD	25.87	21.51
DIVISION	1.93	0.86	FRAC_SD	6.75	4.87	SIMI_RA	26.11	18.75
SIDI	2.35	1.20	SHAPE_CV	7.04	2.89	DCORE_SD	28.62	15.93
SIEI	2.35	1.20	ENN_MN	7.88	3.85	CORE_MN	28.99	19.60
CONTIG_RA	2.35	0.17	CONTIG_CV	8.68	4.92	DCORE_MD	30.83	19.56
CIRCLE_SD	2.52	2.08	CAI_CV	9.47	7.09	PROX_RA	31.25	26.89
CIRCLE_RA	2.88	0.27	LSI	9.65	5.03	AREA_RA	31.65	28.86
PARA_RA	3.43	0.13	SHAPE_SD	9.65	2.81	LPI	31.66	28.91
FRAC_RA	3.56	1.56	PARA_AM	9.76	5.05	DCORE_RA	32.05	18.35
ECON_MN	3.57	1.44	TE	9.85	5.11	DCORE_AM	33.06	17.49
CONNECT	3.57	1.62	ENN_MD	9.86	3.72	PROX_MN	33.10	27.42
SHEI	3.59	2.51	SIMI_CV	9.87	7.02	SIMI_AM	33.25	23.45
SHDI	3.59	2.52	ED	9.87	5.11	CORE_SD	34.51	25.46
CONTIG_MN	3.61	3.76	ENN_SD	10.13	6.14	PROX_AM	34.90	27.58
GYRATE_CV	3.65	3.75	PROX_CV	10.71	8.97	PROX_SD	34.97	29.17
CIRCLE_CV	3.74	4.01	NP	11.86	9.47	SIMI_SD	37.04	34.54
SHAPE_MN	3.80	1.74	PD	11.87	9.47	AREA_AM	37.23	30.32
PARA_CV	3.96	2.46	AREA_MN	11.88	9.48	MESH	37.23	30.32
CONTIG_MD	4.40	3.29	DCORE_CV	12.11	8.85	SPLIT	37.23	30.31
ECON_MD	4.43	0.00	DCAD	12.76	8.70	SIMI_MN	37.36	32.43
IJI	4.45	3.58	NDCA	12.76	8.72	CORE_RA	37.51	29.38
ECON_CV	4.54	3.94	ENN_RA	12.79	7.32	SIMI_MD	40.70	40.74
ENN_AM	4.56	1.12	CORE_CV	12.93	8.19	CORE_AM	54.50	50.61

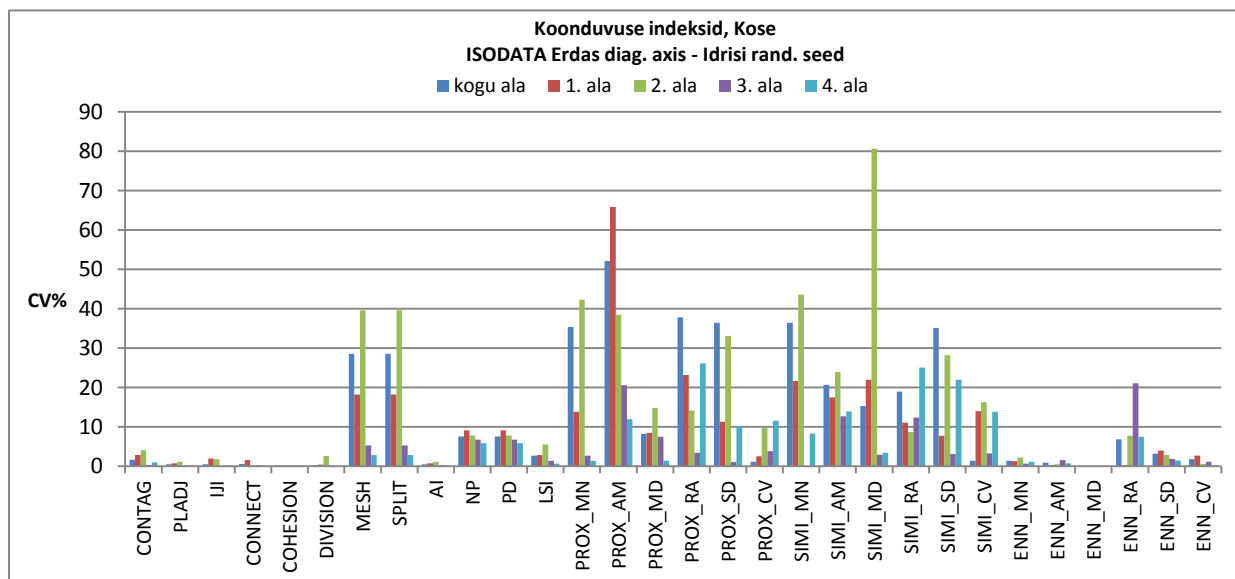
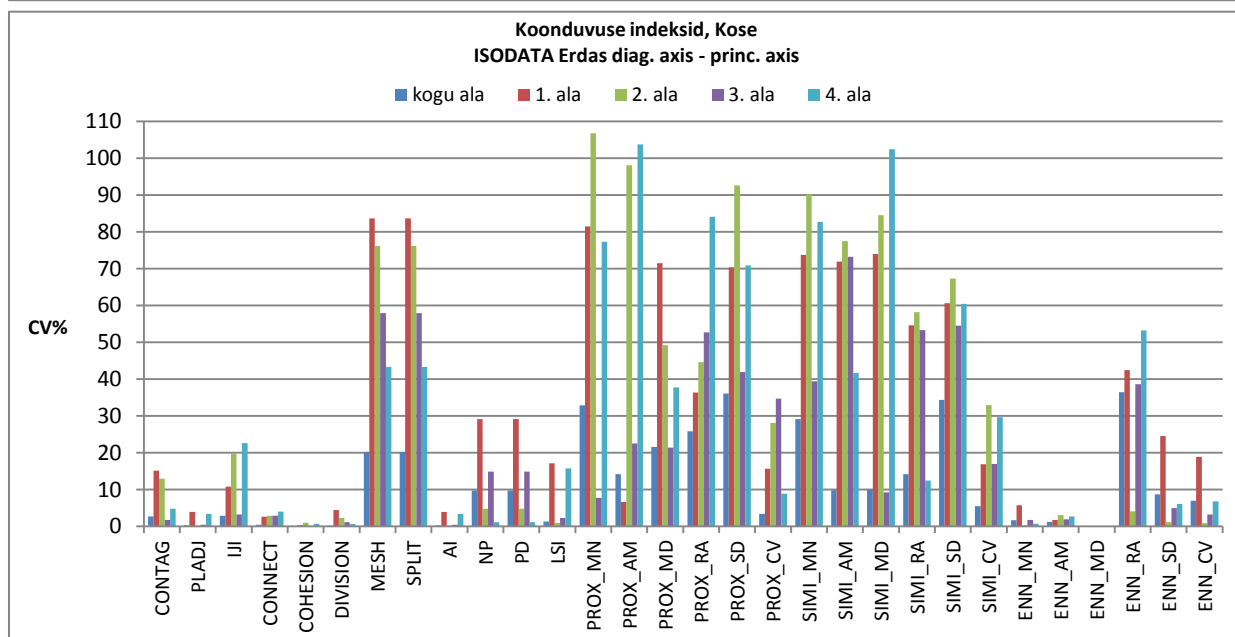
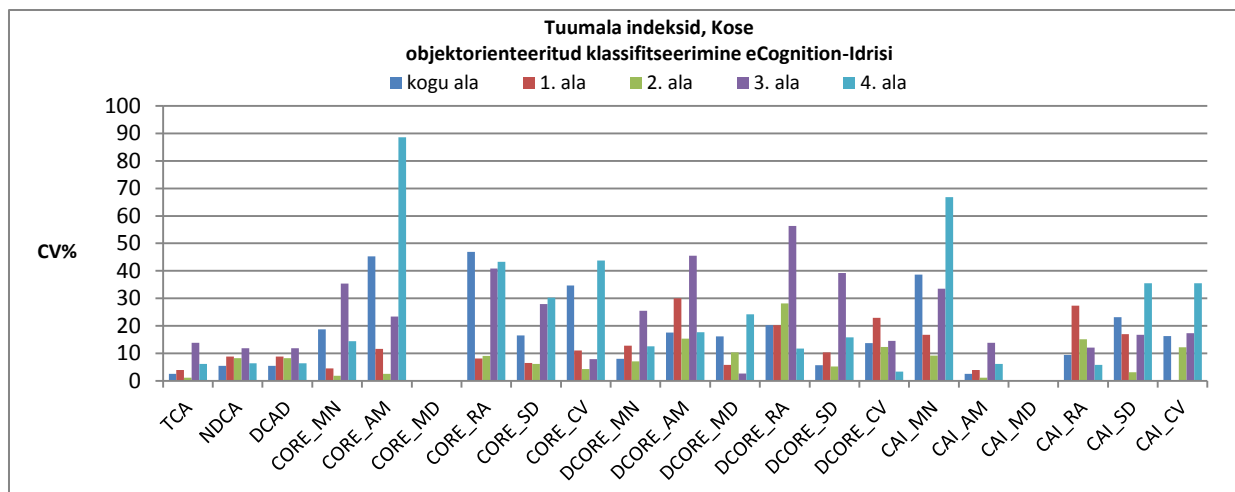
Lisa 4. Indeksite variatsioonikoefitsendid töötlusviiside ning alade lõikes jaotatuna indeksi tüübi alusel (tulemuste osast välja jäänud graafikud)

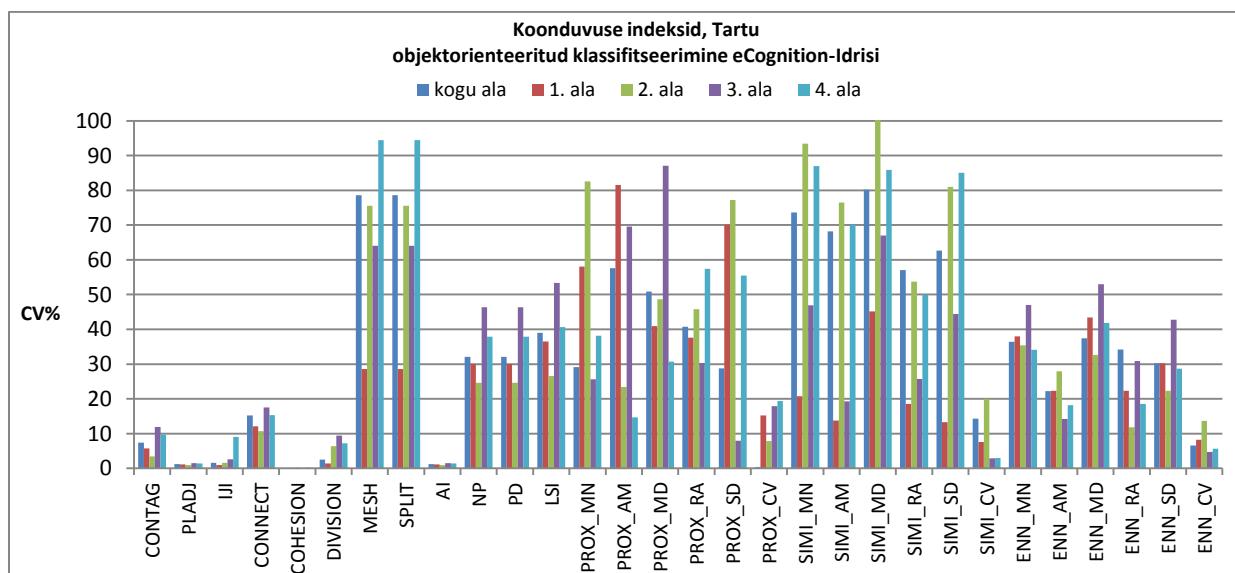
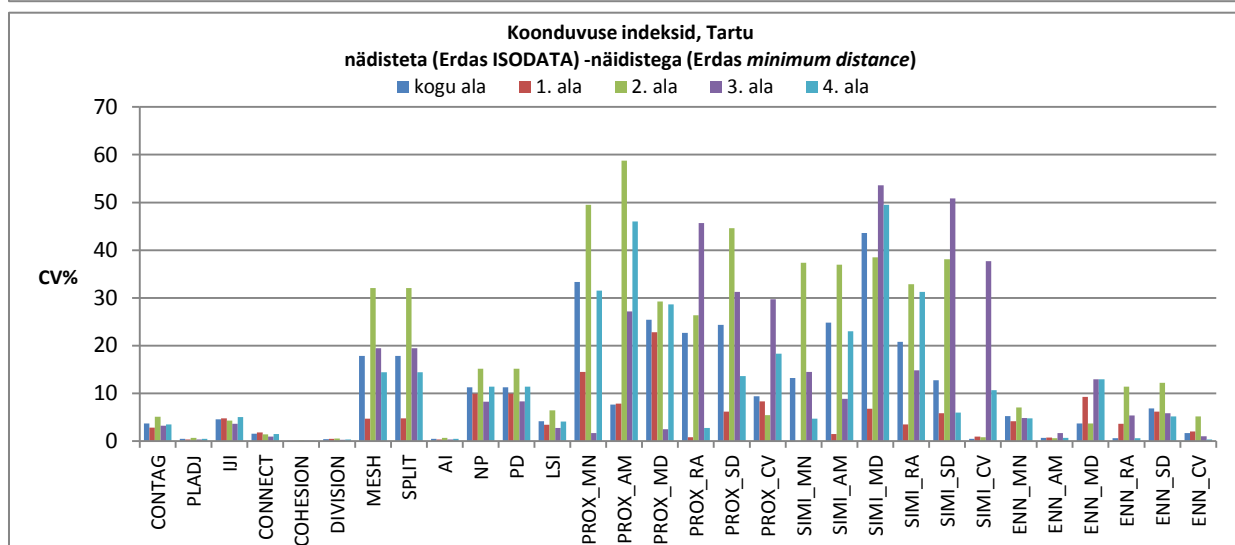
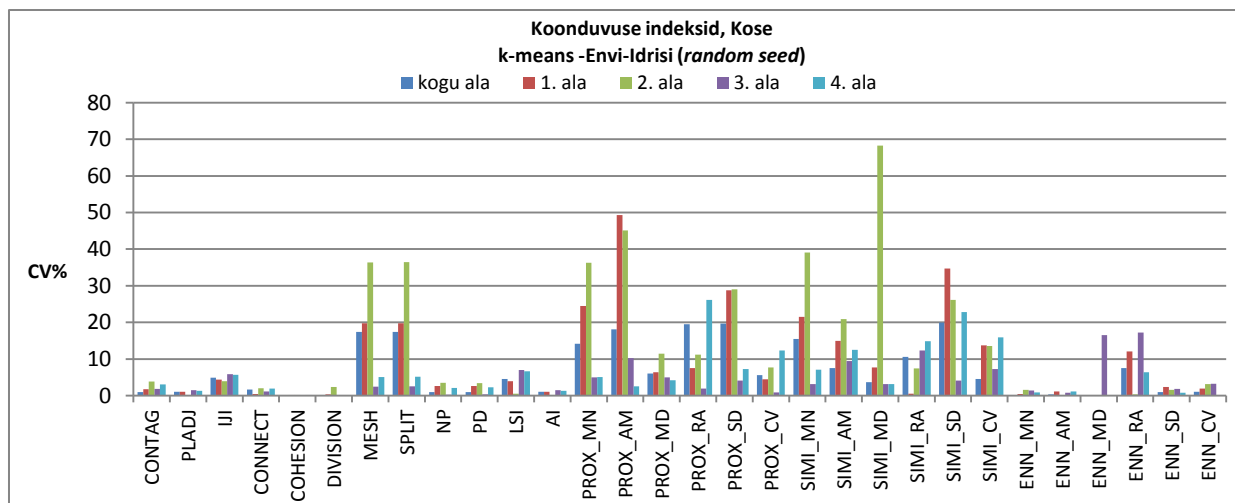


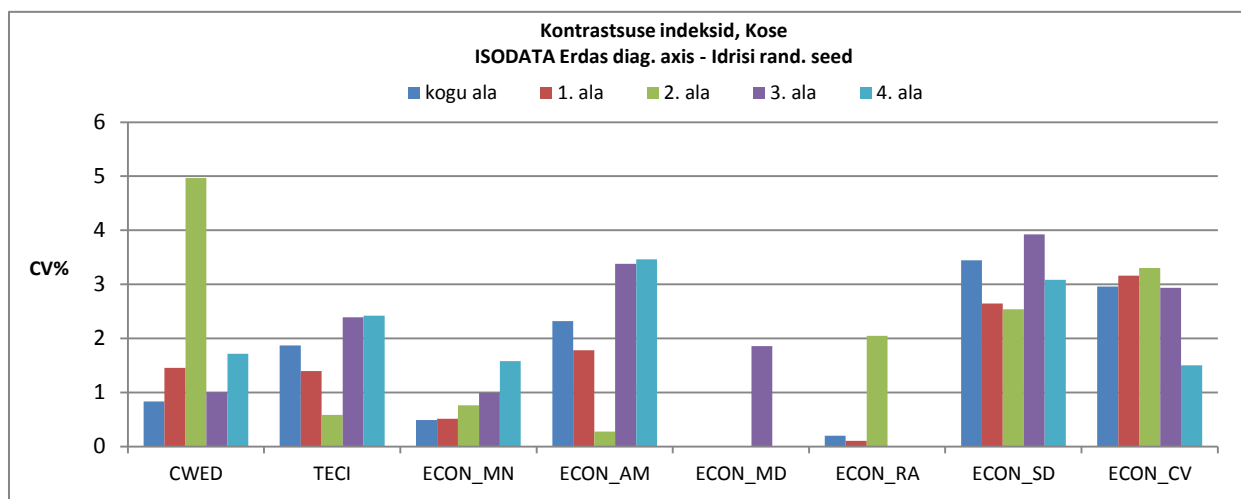
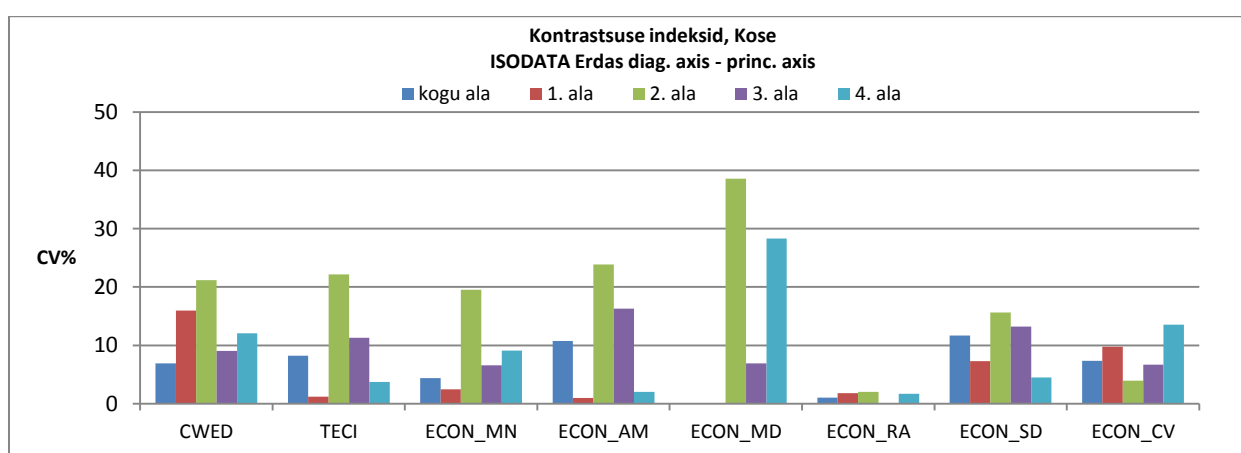
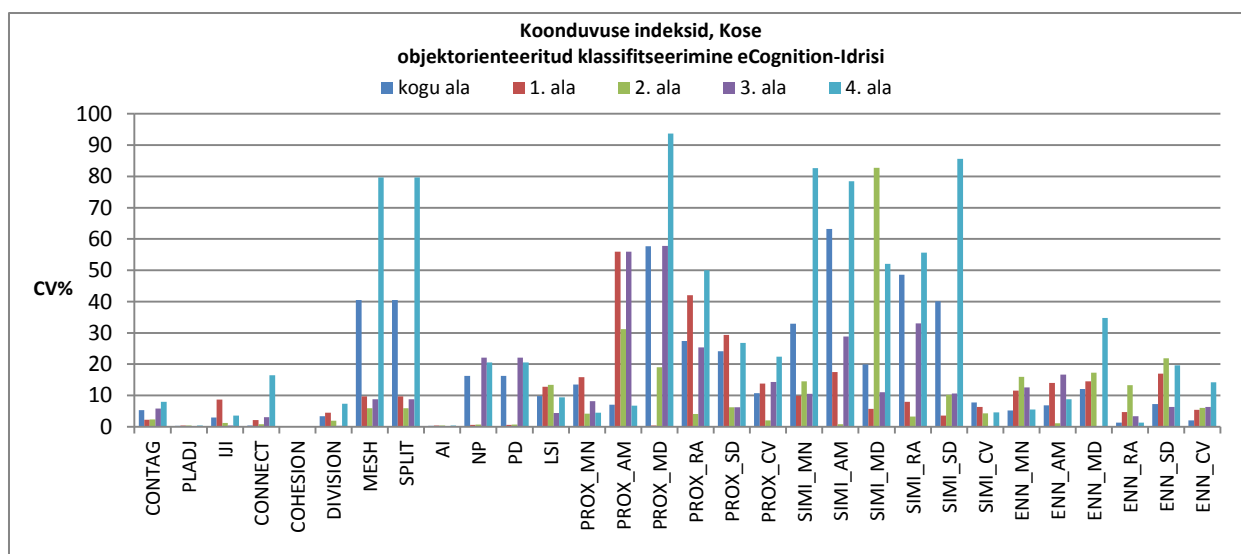


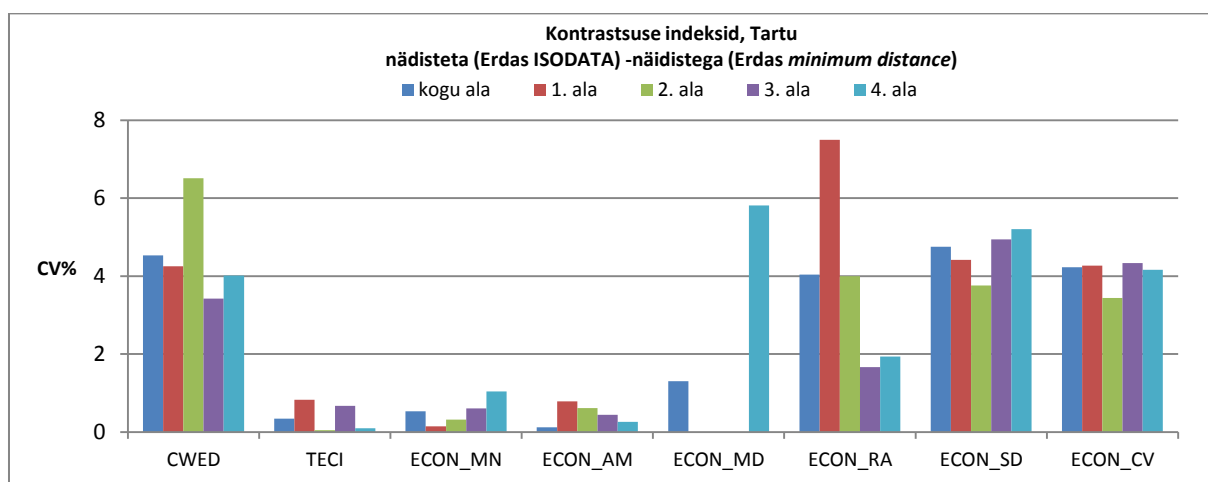
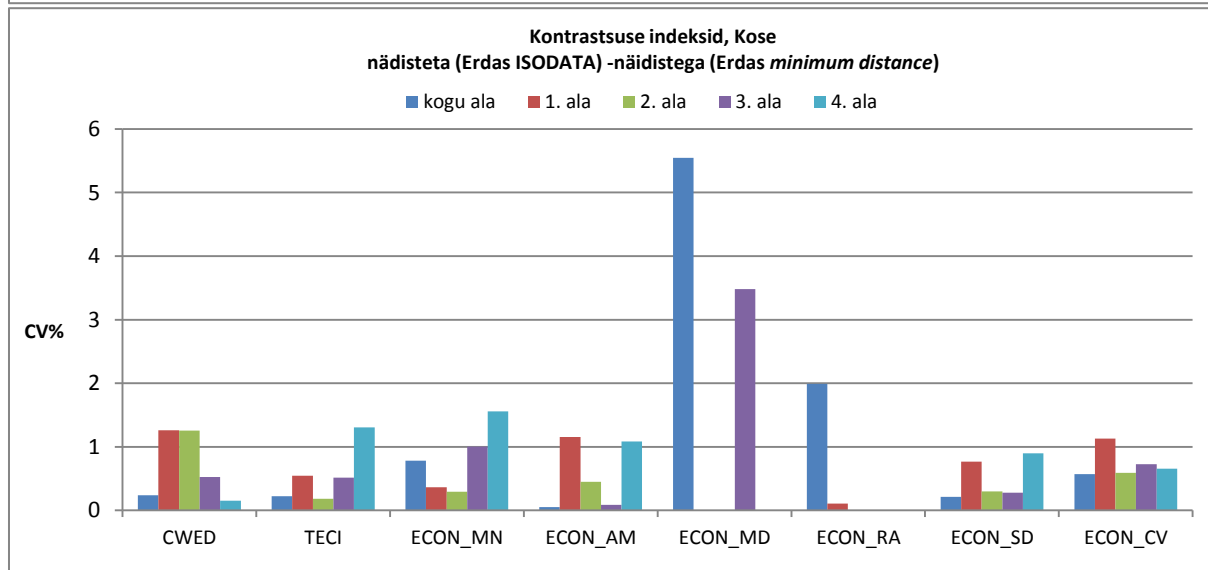
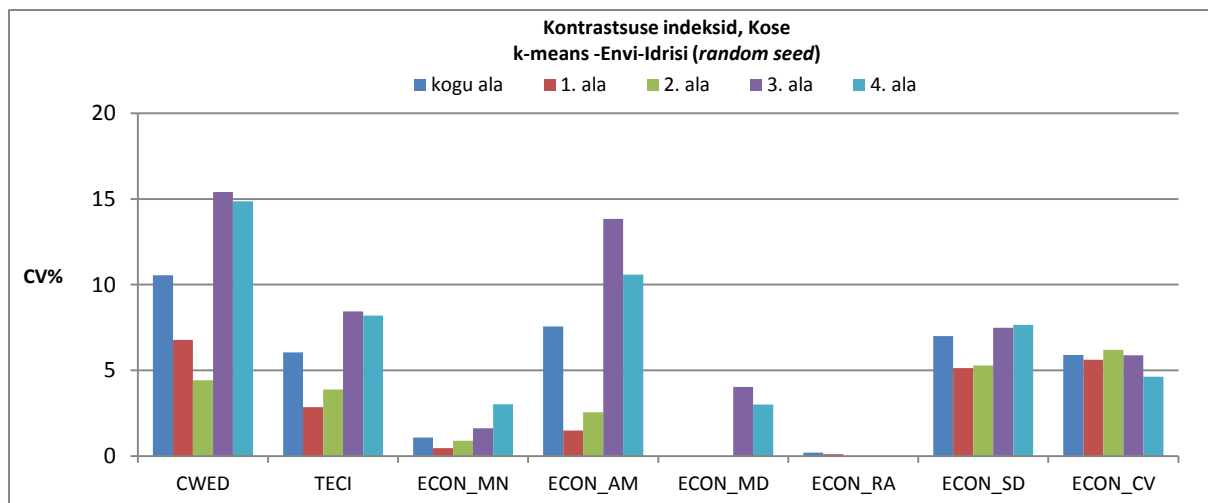


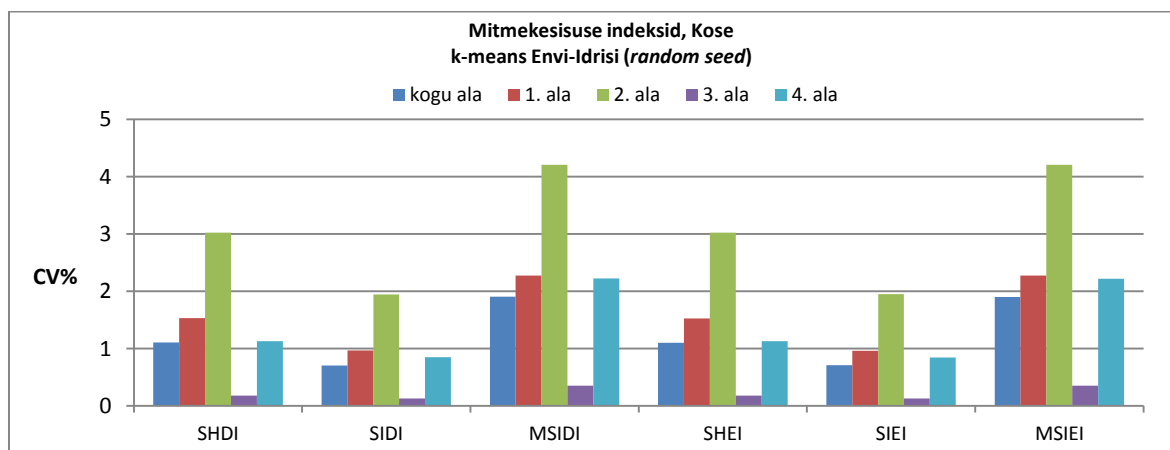
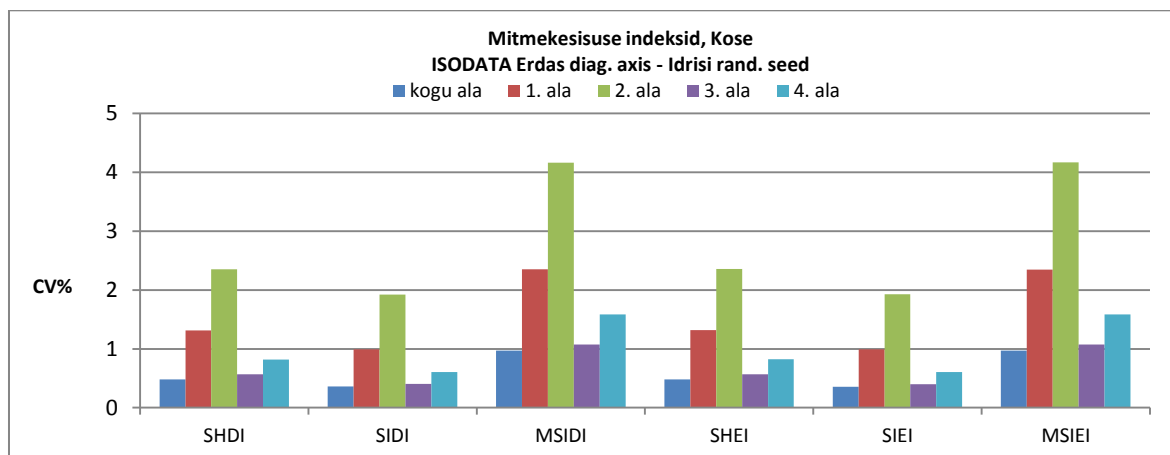
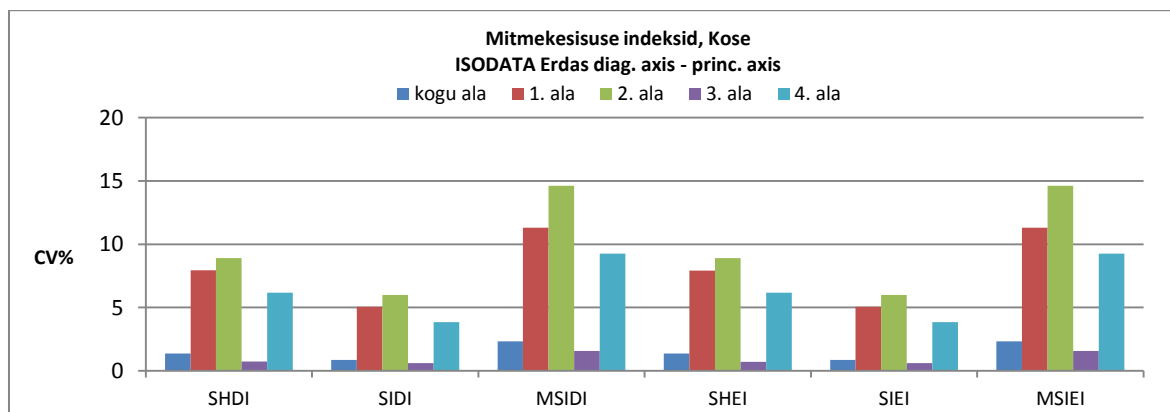
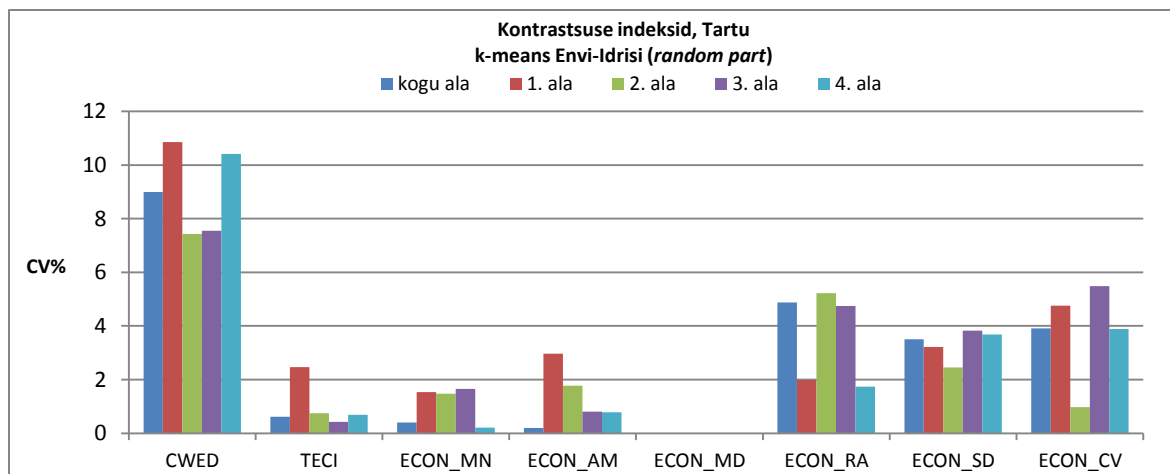


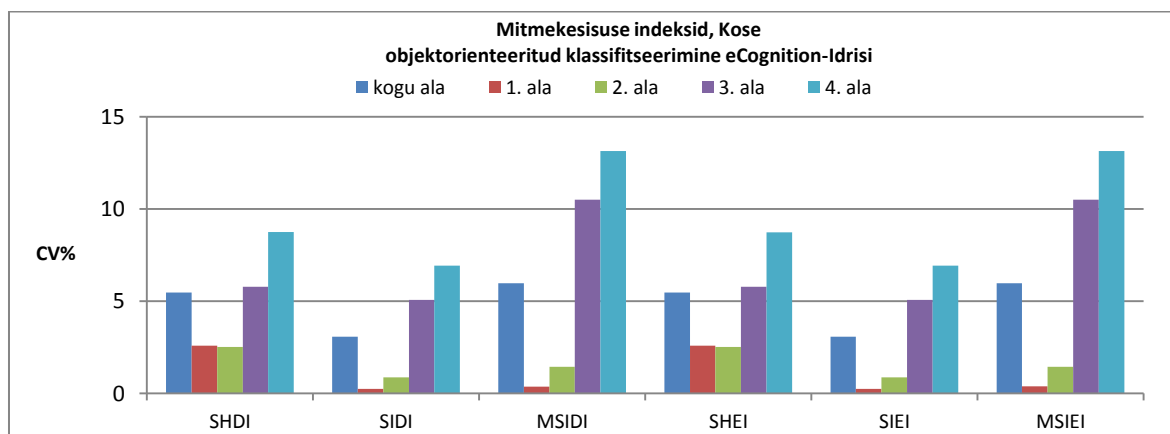
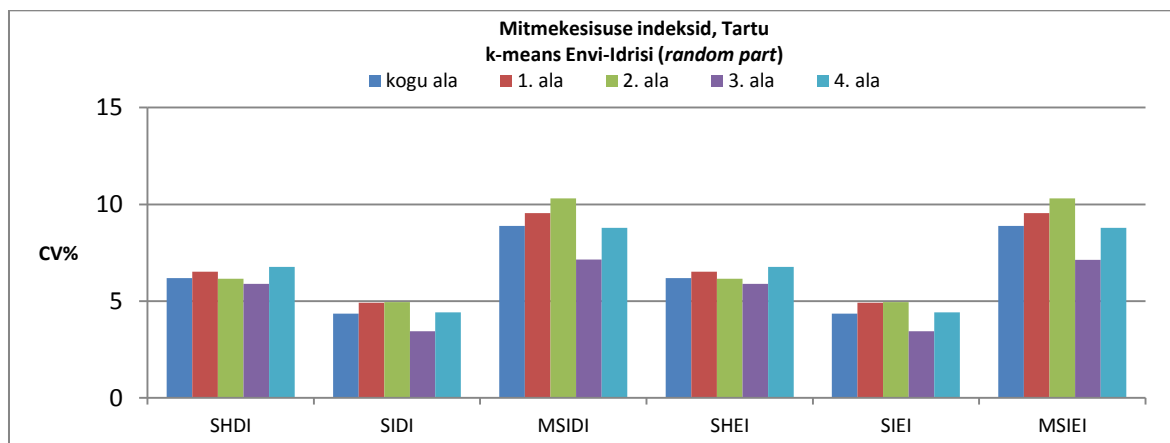
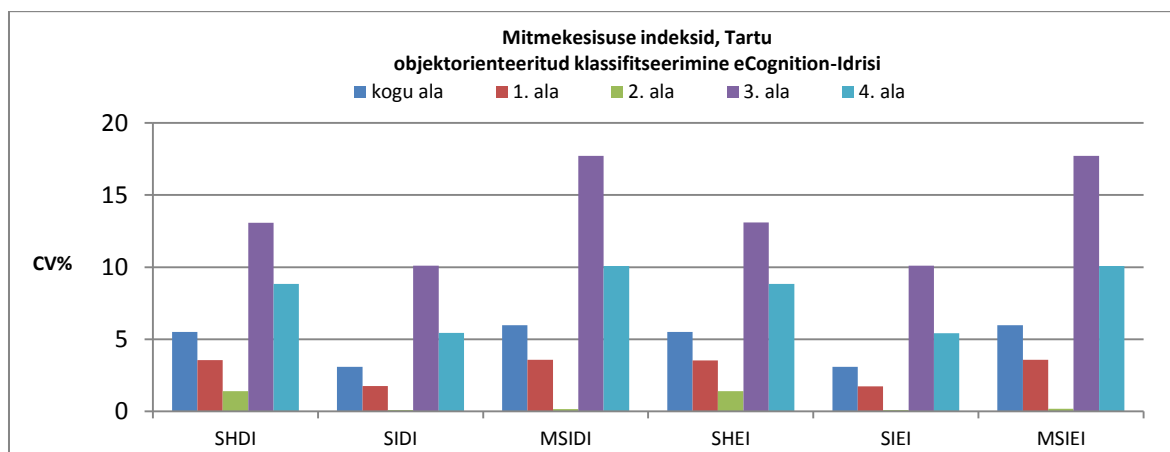
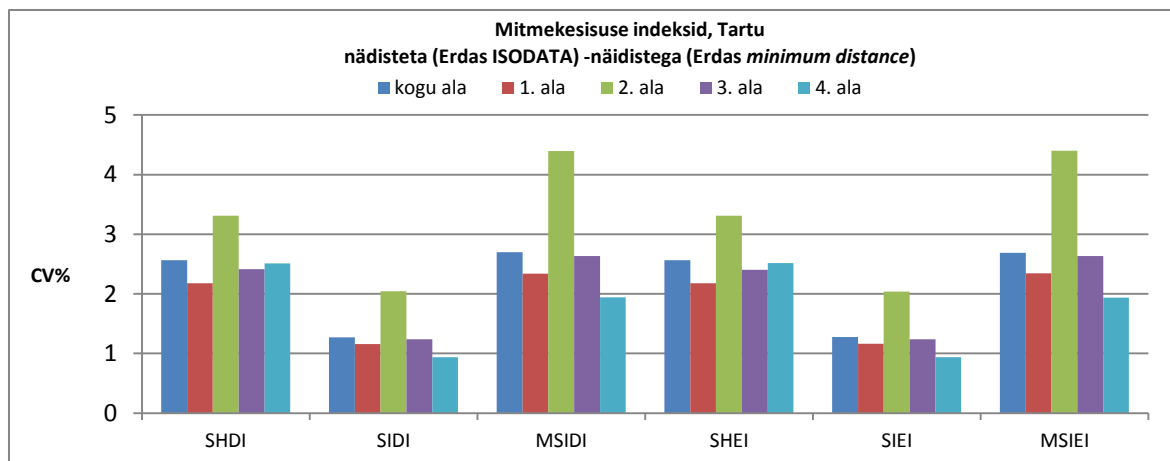












Lihthitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kirke Narusk (*autori nimi*),

annan Tartu Ülikoolile tasuta loa (lihthitsentsi) enda loodud teose

Maastikumustri indikaatorite tundlikkus satelliidipiltide rasterandmestiku töötlustehnikate suhtes (*lõputöö pealkiri*),

mille juhendaja on Tõnu Oja (*juhendaja nimi*),

- 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihthitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 19.05.2014